

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 19, 2003, 16:55:48 ; Search time 2430.79 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-3
Perfect score: 642
Sequence: 1 ATGCCCGGTGCTTCTTC.....TACTCTTGCTCTCTAATAG 642

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.hcg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	642	100.0	642	6	A48665	A48665 Sequence 3
2	642	100.0	642	6	AR157324	AR157324 Sequence
3	642	100.0	642	6	AX452752	AX452752 Sequence
4	642	100.0	642	6	AX685004	AX685004 Sequence
5	628.2	97.9	795	6	A48667	A48667 Sequence 5
6	628.2	97.9	795	6	AR157325	AR157325 Sequence
7	628.2	97.9	795	6	AX452754	AX452754 Sequence
8	628.2	97.9	795	6	AX685006	AX685006 Sequence
9	619.8	96.5	2082	6	A48709	A48709 Sequence 47
10	619.8	96.5	2082	6	AR157350	AR157350 Sequence
11	619.8	96.5	2082	6	AX452796	AX452796 Sequence
12	619.8	96.5	2082	6	AX685048	AX685048 Sequence
13	619.8	96.5	2433	6	A48711	A48711 Sequence 49
14	619.8	96.5	2433	6	AR157351	AR157351 Sequence
15	619.8	96.5	2433	6	AX452798	AX452798 Sequence
16	619.8	96.5	2433	6	AX685050	AX685050 Sequence
17	575	89.6	9379	14	AF207766	AF207766 Hepatitis
18	573.4	89.3	9410	14	HPCKIR2	D50481 Hepatitis C
19	569.6	88.7	9418	14	HCV132996	AJ132996 Hepatitis
20	568.6	88.6	9386	14	AF165056	AF165056 Hepatitis
21	567	88.3	9386	14	AF165055	AF165055 Hepatitis
22	567	88.3	9410	14	HPCKIS2	D50485 Hepatitis C
23	567	88.3	9598	14	AB049101	AB049101 Hepatitis
24	566.4	88.2	3461	6	E07544	E07544 Gene fragment
25	566.4	88.2	3461	6	E09288	E09288 CDNA encodi
26	565.4	88.1	1880	14	HPCSTRJ4	D00832 Hepatitis C
27	565.4	88.1	2540	6	E04260	E04260 CDNA encodi
28	565.4	88.1	2540	6	E04805	E04805 CDNA to 5'
29	565.4	88.1	2540	6	E07391	E07391 CDNA encodi
30	565.4	88.1	3296	14	AB008441	AB008441 Hepatitis
31	565.4	88.1	9033	14	HCV238800	AJ238800 Hepatitis
32	565.4	88.1	9448	14	HPCJ483	D13558 Hepatitis C
33	565.4	88.1	9456	14	HPCBNA	D10934 Hepatitis C
34	565.4	88.1	9605	6	AX739971	AX739971 Sequence
35	565.4	88.1	9605	14	HCV238799	AJ238799 Hepatitis
36	565.4	88.1	11076	6	AX036252	AX036252 Sequence
37	565.4	88.1	11076	6	AX036258	AX036258 Sequence
38	565.4	88.1	11076	6	AX036260	AX036260 Sequence
39	565.4	88.1	11076	6	AX036262	AX036262 Sequence
40	564.8	88.0	1054	14	HPCPEP63	D00689 Hepatitis C
41	563.8	87.8	1539	6	AR027786	AR027786 Sequence
42	563.8	87.8	1863	6	AR027783	AR027783 Sequence
43	563.8	87.8	3296	14	AB008442	AB008442 Hepatitis
44	563.8	87.8	3296	14	AB008443	AB008443 Hepatitis
45	563.8	87.8	9379	14	HCVPOLYP	AJ000009 Hepatitis

ALIGNMENTS

RESULT 1
A48665
LOCUS A48665 Sequence 3 from Patent WO9604385. 642 bp DNA linear PAT 07-MAR-1997
DEFINITION A48665
ACCESSION A48665
VERSION A48665.1 GI:2302378
KEYWORDS
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 642)
AUTHORS Maertens G., Rosman F., De M.G. and Buyse M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 3 15-FEB-1996;

ACCESSION AX452752

RESULT 2
AR157324

VERSION	AX452752.1	GI:21712437
KEYWORDS	Hepatitis C virus	
SOURCE	Hepatitis C virus	
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.	
REFERENCE	1	
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buysse,M.A.	
TITLE	Recombinant vectors for producing hcv envelope proteins	
JOURNAL	Patent: EP 1211315-A 3 05-JUN-2002;	
INNOGENETICS	Innogenetics N.V. (BE)	
FEATURES	Location/Qualifiers	
source	1..642	
	/organism="Hepatitis C virus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:11103"	
CDS	1.._639	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD38066.1"	
	/db_xref="GI:21712438"	
	/translation="MPCGSPSIFLLALLSLCTIPASAVEVRNYSGMVHTNDSCNSIIVYEADMIMTGPCVPCVRNNSSRCWALPTILARNASVPPTTIIRHVDDLVGAAALCSAMYGDLCGVSVLFSYFTTISPRRHETVDQNCISIYPGHITGHRNADMMNMNSPITALVSQLLRIPOVVAVDMVAGHWGLAGLYSYMGVNKVLIVMLLFAL"	
mat_peptide	1..636	
BASE COUNT	109 a 195 c 176 g 162 t	
ORIGIN		
Query Match	100.0%; Score 642; DB 6; Length 642;	
Best Local Similarity	100.0%; Pred. No. 1.le-150;	
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 ATGCCCGGTGGCTTTCTCTATCTTCCTCTTGCTTACTGTCTCCTGTGACCAATTCCA	60
Dd	1 ATGCCCGGTGGCTTTCTCTATCTTCCTCTTGCTTACTGTCTCCTGTGACCAATTCCA	60
Qy	61 GTTTCGGCTTATGAGGTGCACACGTGTCCGGATGTACCATGTACAGACGACTGTCTC	120
Dd	61 GTTTCGGCTTATGAGGTGGCAAAGTGTCCGGATGTACCATGTACAGACGACTGTCTC	120
Qy	121 AACTCAAGCATTTGTATGAGCAGCGGACAATGATCATGACACCCCCGGGTGGTGCCTC	180
Dd	121 AACTCAAGCATTTGTATGAGCAGCGGACAATGATCATGACACCCCCGGGTGGTGCCTC	180
Qy	181 TGCCTTCGGAGAACAACTCTTCCTGGCTGTGGTAGGCTCACCCCACGCTCGCAGCT	240
Dd	181 TGCCTTCGGAGAACAACTCTTCCTGGCTGTGGTAGGCTCACCCCACGCTCGCAGCT	240
Qy	241 AGGAACGCACGCTCCCAACACACAATACGACGCCACGTGATTTGCTGTTGGGGCG	300
Dd	241 AGGAACGCACGCTCCCAACACACAATACGACGCCACGTGATTTGCTGTTGGGGCG	300
Qy	301 GCTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGCTCCTCGTCTCC	360
Dd	301 GCTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGCTCCTCGTCTCC	360
Qy	361 CAGCTGTTACCATCTCGCTCCGGCATGAGAAGGTGACAGGACTGCAATGTCTCAATC	420
Dd	361 CAGCTGTTACCATCTCGCTCCGGCATGAGAAGGTGACAGGACTGCAATGTCTCAATC	420
Qy	421 TATCCCGCCACATAAAGGTCACCGTATGCTTGGGATATGATGTAAGTGGTTCGCT	480
Dd	421 TATCCCGCCACATAAAGGTCACCGTATGCTTGGGATATGATGTAAGTGGTTCGCT	480
Qy	481 ACAACGGCCCTTGGTGTATCGCAGCTGCTCCGGATCCCAAGCTGCTGCGACATGGTG	540
Dd	481 ACAACGGCCCTTGGTGTATCGCAGCTGCTCCGGATCCCAAGCTGCTGCGACATGGTG	540
Qy	541 CGCGGGGCCCATNTGGGAGTCTCGCGGCTCGCCCTACTATTCCATGGTGGGAACTGG	600
Dd	541 CGCGGGGCCCATNTGGGAGTCTCGCGGCTCGCCCTACTATTCCATGGTGGGAACTGG	600

Qy	601	GCTAAGGTTTGTGATGTGATGCTACTCTTTGCTCTCTTAATAG	642	linear	PAT 29-MAR-2003
Db	601	GCTAAGGTTTGTGATGTGATGCTACTCTTTGCTCTCTTAATAG	642		
RESULT 4					
AX685004					
LOCUS	AX685004	642 bp	DNA		
DEFINITION	Sequence 3 from Patent WO02055548.				
ACCESSION	AX685004				
VERSION	AX685004.1	GI:29371409			
KEYWORDS					
SOURCE	Hepatitis C virus				
ORGANISM	Hepatitis C virus				
REFERENCE	1				
AUTHORS	Maertens G., Boorman F. and Buyse, M.A.				
TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use				
JOURNAL	Patent: WO 02055548-A 3 18-JUL-2002;				
FEATURES	INNOGENETICS N.V. (BE)				
source	Location/Qualifiers				
	1..642				
	/organism="Hepatitis C virus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:11103"				
CDS	1..639				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD06519.1"				
	/db_xref="GI:29371410"				
	/translation="MPCGSPSIFLLALLSLCTIPASAYEVRNVSGMVHVTNDCNSISIVYAADIMHTPGCPVCRNENSSRCWVALTPTLAARNASVPPTTIRRHVDLLVGAALVCSAMYVGDLCGSVFLVSQFTISPRRHETVQCNCISYPGHITGHRMADWMMWNSPFTALVSGOLLRIPOAVVDVMVAGHWGLAGLAYVSMVGNKVLIVMLLPAL"				
mat_peptide	1..635				
	/product="unnamed"				
BASE COUNT	109 a 195 c 176 g 162 t				
ORIGIN					
Query Match	100.0%;	Score 642;	DB 6;	Length 642;	
Best Local Similarity	100.0%;	Pred. No. 1.1e-150;			
Matches 642;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGCCCGGTTGCTCTTTCTCTATCTTCTCTCTTGCGTTTACTGTCTCTGTGACCAATCCCA	60		
Db	1	ATGCCCGGTTGCTCTTTCTCTATCTTCTCTTGCGTTTACTGTCTCTGTGACCAATCCCA	60		
Qy	61	GCTTCCGGTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTACGACGACTGCTCC	120		
Db	61	GCTTCCGGTTATGAGGTGCGCAACGTTGTCGGGATGTACCATGTACGACGACTGCTCC	120		
Qy	121	AACCTCAAGCATGTGTATGAGGACGGACATGATATGCACACCCCGGGTGGTGCGCC	180		
Db	121	AACCTCAAGCATGTGTATGAGGACGGACATGATATGCACACCCCGGGTGGTGCGCC	180		
Qy	181	TGCGTTTCGGGAGAACAACTCTTCGCCGTGCTGGGTAGCGGTCAACCCACGCTCGCAGCT	240		
Db	181	TGCGTTTCGGGAGAACAACTCTTCGCCGTGCTGGGTAGCGGTCAACCCACGCTCGCAGCT	240		
Qy	241	AGGAACGGCAGCGTCCCGCAACGACAAATACGACGCCACGTCGATTTGCTGTTGGGGCG	300		
Db	241	AGGAACGGCAGCGTCCCGCAACGACAAATACGACGCCACGTCGATTTGCTGTTGGGGCG	300		
Qy	301	GCTGCTCTGTGTTCCGGTATGTAGTGGGGATCTCTGCGGATCTGTTCTCTGCTCTCC	360		
Db	301	GCTGCTCTGTGTTCCGGTATGTAGTGGGGATCTCTGCGGATCTGTTCTCTGCTCTCC	360		
Qy	361	CAGCTGTTTACCACTCTCGCCTCGCCGGCATGAGACGGGTGAGGACTGCAATTTGCTCAATC	420		
Db	361	CAGCTGTTTACCACTCTCGCCTCGCCGGCATGAGACGGGTGAGGACTGCAATTTGCTCAATC	420		

Db	335	GGGTTCCGGAGAAACACTCTTCCGGCTGCTGGGTAGCGCTCA	CCCCACGCTCGCAGCTA	394
Qy	242	GGAAAGCCAGCGTCCCAACACGACAAATACAGACCCACGTCGA <td>TTTGGTCTCGTTGGGGCGG</td> <td>301</td>	TTTGGTCTCGTTGGGGCGG	301
Db	395	GGAAAGCCAGCGTCCCAACACGACAAATACAGACCCACGTCGA <td>TTTGGTCTCGTTGGGGCGG</td> <td>454</td>	TTTGGTCTCGTTGGGGCGG	454
Qy	302	CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCT	GTCTTCTCTCGTCTCCC	361
Db	455	CTGCTTCTGTTCCGCTATGTACGTGGGGGACCTCTGGGATCT	GTCTTCTCTCGTCTCCC	514
Qy	362	AGCTGTTACACCTCTCGCTCCGCGGATGAGACGTCGAGACT	GCAAATGCTCAATCT	421
Db	515	AGCTGTTACACCTCTCGCTCCGCGGATGAGACGTCGAGACT	GCAAATGCTCAATCT	574
Qy	422	ATCCCGGCCACATAACAGAGTCAACGATGCTGGGATATGATG	ATGATGATGATGATGATGATGATGATGATGATGATGATGATG	481
Db	575	ATCCCGGCCACATAACAGAGTCAACGATGCTGGGATATGATG	ATGATGATGATGATGATGATGATGATGATGATGATGATGATG	634
Qy	482	CAACGCGCTGCTGGTATCGCAGCTGCTCCGGATCCCAAGCT	TGTCGTGGACATGCTGGGATG	541
Db	635	CAACGCGCTGCTGGTATCGCAGCTGCTCCGGATCCCAAGCT	TGTCGTGGACATGCTGGGATG	694
Qy	542	CGGGGGCCCAATGGGGAGTCTGGGGGCTCGCTACTATTCAT	TCCATGCTGGGGAACCTGGG	601
Db	695	CGGGGGCCCAATGGGGAGTCTGGGGGCTCGCTACTATTCAT	TCCATGCTGGGGAACCTGGG	754
Qy	602	CTAAGGTTTCATGTCATGCTACTCTTTGCTCTCTAATAG	642	
Db	755	CTAAGGTTTCATGTCATGCTACTCTTTGCTCTCTAATAG	795	
RESULT 6				
LOCUS	AR157325	Sequence 5 from patent US 6245503.	795 bp	DNA linear PAT 17-OCT-2001
DEFINITION	AR157325	Sequence 5 from patent US 6245503.		
ACCESSION	AR157325.1	GI:16218258		
VERSION	AR157325.1	GI:16218258		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 795)			
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.			
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use			
JOURNAL	Patent: US 6245503-A 5 12-JUN-2001;			
FEATURES	Location/Qualifiers			
source	1..795			
BASE COUNT	130 a 240 c 231 g 194 t			
ORIGIN				
Query Match	97.9%;	Score 628.2;	DB 6;	Length 795;
Best Local Similarity	98.8%;	Pred. No. 3.2e-147;		
Matches	633;	Conservative	0;	Mismatches 8; Indels 0; Gaps 0;
Qy	2	TGCCCGGTTGCTCTTTCTCTATCTTCCTCTTGGCTTTACTGTCTGTCTGACCAATTCAG	61	
Db	155	TGCCCGGTTGCTCTTTCTCTATCTTCCTCTTGGCTTTGCTGTCTGTCTGACCGTTTCAG	214	
Qy	62	CTTCCGCTTATGAGTTCGCGCAAGTTCGCGGATGTACATGTGTACAGAACGACTGCTCCA	121	
Db	215	CTTCCGCTTATGAGTTCGCGCAAGTTCGCGGATGTACATGTGTACAGAACGACTGCTCCA	274	
Qy	122	ACTCAAGCATTGTGTATGAGGACGCGGACATGATCATGTACACACCCCGGGTGCCTGCCT	181	
Db	275	ACTCAAGCATTGTGTATGAGGACGCGGACATGATCATGTACACACCCCGGGTGCCTGCCT	334	
Qy	142	GGTTCGGGAGAACACTCTTCGCTGTGGTGTAGCGTCTACCCGCCACGCTCGCAGCTA	241	
Db	335	GGTTCGGGAGAACACTCTTCGCTGTGGTGTAGCGTCTACCCGCCACGCTCGCAGCTA	394	

	Qy	421	TATCCGGCCACATAAAGGTCACCGTATGCCTTGGGATATGATGATGAACACTGGTCGCCT	480
	Db	421	TATCCGGCCACATAAAGGTCACCGTATGCCTTGGGATATGATGATGAACACTGGTCGCCT	480
	Qy	481	ACAACGGGCCCTGGTGATCGCAGCTGCTCCGGATCCCACAAGTGTCTGTGGACATGGTG	540
	Db	481	ACAACGGGCCCTGGTGATCGCAGCTGCTCCGGATCCCACAAGTGTCTGTGGACATGGTG	540
	Qy	541	GCGGGGGCCCATTTGGGGAGTCTCTGGGGCTCGCTACTATTCCATGGTGGGAACTGG	600
	Db	541	GCGGGGGCCCAITTTGGGGAGTCTCTGGGGCTCGCTACTATTCCATGGTGGGAACTGG	600
	Qy	601	GCTTAAGGTTTTGATCTGATGCTACTCTTTGCTCTCTAATAG	642
	Db	601	GCTTAAGGTTTTGATGTTGATGCTACTCTTTGCTCTCTAATAG	642
	RESULT 5			
	A48667	A48667	795 bp	DNA linear PAT 07-MAR-1997
	LOCUS	Sequence 5 from Patent WO9604385.		
	DEFINITION	A48667		
	ACCESSION	A48667.1	GI:2302380	
	VERSION			
	KEYWORDS	unidentified		
	SOURCE	unclassified.		
	ORGANISM	unclassified.		
	REFERENCE	1 (bases 1 to 795)		
	AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buyse,M.		
	TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
	JOURNAL	Patent: WO 9604385-A 5 15-FEB-1996;		
		INNOGENETICS NV (BE)		
	COMMENT	Other publication CA 2172273 960215		
		Other publication AU 3382495 960304.		
	FEATURES	Location/Qualifiers		
	source	1..795		
		/organism="unidentified"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:32644"		
	CDS	1..792		
		/note="unnamed protein product"		
		/codon_start=1		
		/protein_id="CAA03129.1"		
		/db_xref="GI:2302381"		
		/translation="MLGRVIDTLTCGFADLVGYIPLVGLPGGNARALAHGVRLVDG VNATGNLPGCSFSLFLIALASCLTPASAYEVRNVSMGHVTNDSCNSIIVYEAADM TMHTPCVCPIRENNSSRCWALTPTLAARNASVPTTIIRHVDLLVGLAAAFCSAMTV GDLCSFVLVSLEFTISPRRHETQDCNCISYPGHIIGHRMWMNMNWSPTTVALVPS QLULRIPQAVDMVACAHGWLAGLAYMSVGNWAKVLIVMLLPAP"		
	mat_peptide	1..789		
		/product="unnamed"		
	BASE COUNT	130 a 240 c 231 g	194 t	
	ORIGIN			
	Query Match	97.9%;	Score 628.2; DB 6;	Length 795;
	Best Local Similarity	98.8%;	Prod.No.3.2e-147;	
	Matches	633; Conservative	0; Mismatches 8;	Indels 0; Gaps 0;
	Qy	2	TGCCCGGTGCTCTTTCTCTATCTCTCTCTGCTTTTACTGTCTGTCTGACCATTCCAG	61
	Db	155	TGCCCGGTGCTCTTTCTCTATCTCTCTGCTTTTACTGTCTGTCTGACCATTCCAG	214
	Qy	62	CTTCCGCTTATGAGGTGCGCAACGTGTCCGGATGATACCATGTGCACGAACGACTGCTCCA	121
	Db	215	CTTCCGCTTATGAAAGTGGCAAGCTGTCCGGATGATACCATGTGCACGAACGACTGCTCCA	274
	Qy	122	ACTCAAGCATTTGTATGAGGACGGACATGATCATGCACACCCCGGGTGGTGGCCCT	181
	Db	275	ACTCAAGCATTTGTATGAGGACGGACGACATGATCATGCACACCCCGGGTGGTGGCCCT	334
	Qy	182	CGCTTCGGGAGAACAACTCTTCCTCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTA	241

mat_peptide		1..789	source		1..2082
BASE COUNT		130 a 240 c 231 g 194 t	/organism="unidentified"		/mol_type="genomic DNA"
ORIGIN			/db_xref="taxon:32644"		
Query Match		97.9%; Score 628.2; DB 6; Length 795;	/note="unnamed protein product; Protein sequence is in conflict with the conceptual translation"		
Best Local Similarity		98.8%; Pred. No. 3.2e-147;	/codon_start=1		
Matches 633; Conservative		0; Mismatches 8; Indels 0; Gaps 0;	/protein_id="CAA03146.1"		
			/db_xref="GI:2302423"		
			/translation="NLGKVIDTLTCGADLVGVIPLVGAPLGGARALAHGVRVLEDEG VNYATGNLPGCSF8IFILLALLSLTVPASATEVRNVSVMHVHTDNCSSSIYEAADIMHTPGCVPCVRNENSRCSWALPTTLAARNASVPTTIRRHVDLLVGAARFCAAMVY QDLGSLVLSQLEFTISPRHETVQDCNSIYPGHTIHRMAWMMNWSPTIALVVS GLLRTPQAVDMVGAHWVLGLAYLSMVGNWAKLVLMLLFAGVDGHTFRVSGGAA SDDTGLVLSFSPGSAQIOLVNTNGSHINRTALNCNDSLOTGFFAALFYKHFNSSG CPERIASCSDIKFAQGWGPLYTPEBNSDORPYCHVAPRPGIVPASOVCGPVYCF TPSPVAVTDDTRFGVPTYNWANDSDVLI LNNTBPPBGNWFGCTWMNGTGTCTCGGP FCNTGGAGNNTLTCTDCTFRKHPEATYARCGSGFWLTPRCMWHYPYRLMWHYECTVNET IPKVMYVGVGVEHRFEAANCNWRGERCDLEDRSELSPLLSLTEMQILLFCSFTTLP ALSTGLIHQNIVDVQYLVGSAVSLVKWEYVLLFLLLADARICACLMWMLLI AQAEALENLVNLNAAAVAGAHGTLFLVFFCAAAYIKGRLVPGAAYAFYGVWPLLLI LLAUPPRAYA"		
			1..2076		
			/product="unnamed"		
mat_peptide		366 a 634 c 600 g 482 t			
BASE COUNT		366 a 634 c 600 g 482 t			
ORIGIN					
Query Match		96.5%; Score 619.8; DB 6; Length 2082;			
Best Local Similarity		98.9%; Pred. No. 4.2e-145;			
Matches 624; Conservative		0; Mismatches 7; Indels 0; Gaps 0;			
QY	2	TGCCCCGTTGCTCTTTCTCTATCTTCTCTTGGCTTTACTGTCTCTGTCTGACCATCCAG	61		
DB	155	TGCCCCGTTGCTCTTTCTCTATCTTCTCTTGGCTTTGCTGTCTGTCTGACCGTCCAG	214		
QY	62	CTTCCCGCTTATGAGGTGCGCAACGTGTCGGGATGTACCATGTTCAGAACGACTGTCTCA	121		
DB	215	CTTCCCGCTTATGAGGTGCGCAACGTGTCGGGATGTACCATGTTCAGAACGACTGTCTCA	274		
QY	122	ACTCAAGCATTTGTATGAGGACGCGACATGATCATATGCACACCCCGGGTGGTCCCT	181		
DB	275	ACTCAAGCATTTGTATGAGGACGCGACATGATCATATGCACACCCCGGGTGGTCCCT	334		
QY	182	GCCTTCGGGAGAACAACTCTTCCCGCTGTGGTAGCGCTCACCCCGCGCTCGAGCTA	241		
DB	335	GCCTTCGGGAGAACAACTCTTCCCGCTGTGGTAGCGCTCACCCCGCGCTCGAGCTA	394		
QY	242	GGAACGCGCGTCCCGACCAACATAGACGCGCACCGTCCGATTTGCTGTTGGGCGG	301		
DB	395	GGAACGCGCGTCCCGACCAACATAGACGCGCACCGTCCGATTTGCTGTTGGGCGG	454		
QY	302	CTGCTCTCTGTCGGCTATGTACGTGGGGATCTCTGGGATCTGTCTCTCTCTCC	361		
DB	455	CTGCTCTCTGTCGGCTATGTACGTGGGGATCTCTGGGATCTGTCTCTCTCTCC	514		
QY	362	AGCTGTTACCAATCTCCCTCGCGCATGAGCGGTGACAGGACTGCAATTTGCTCAATCT	421		
DB	515	AGCTGTTACCAATCTCCCTCGCGCATGAGCGGTGACAGGACTGCAATTTGCTCAATCT	574		
QY	422	ATCCCGGCGACATAACAGGTCAACCGTATGGCTTGGGATATGATGATGAATGTGTCGCTA	481		
DB	575	ATCCCGGCGACATAACAGGTCAACCGTATGGCTTGGGATATGATGATGAATGTGTCGCTA	634		
QY	482	CAACGCGCTCTGGTATCGCAGCTGTCTCGGATCCCAAGCTGTCTGACATGTGG	541		
DB	635	CAACGCGCTCTGGTATCGCAGCTGTCTCGGATCCCAAGCTGTCTGACATGTGG	694		
QY	542	CGGGGCGCATTTGGGAGTCTCTGGCGGCTCTCGCTACTATTCCATGTTGGGAACTGGG	601		
DB	695	CGGGGCGCATTTGGGAGTCTCTGGCGGCTCTCGCTACTATTCCATGTTGGGAACTGGG	754		
QY	602	CTAAGGTTTTGATTTGATGATGCTACTCTTTTGC	632		

RESULT 9								
LOCUS	A48709	2082 bp	DNA	linear	PAT 07-MAR-1997			
DEFINITION	Sequence 47 from Patent WO9604385.							
ACCESSION	A48709							
VERSION	A48709.1	GI:2302422						
KEYWORDS	unidentified							
SOURCE	unclassified.							
ORGANISM	unclassified							
REFERENCE	1 (bases 1 to 2082)							
AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buysse,M.							
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE							
JOURNAL	Patent: WO 9604385-A 47 15-FEB-1996;							
COMMENT	INNOGENETICS NV (BE)							
	Other publication CA 2172273 960215							
	Other publication AU 3382495 960304.							
FEATURES	Location/Qualifiers							

GVNYATGNLPGCSFSLFLLALLSCLTPVAPASAYEVNRNVGMHYHVNDCSSSIVYEAD
MINHTPGCPVPCVRENNSSRCWALVITPLAARNASVPTTIRRHVDLLVGAAFCSAMY
VGDLCGSLFVLSQIFTSIPRHEVQDCNCSYIPGHTGHRMAWDMWMSPTALVV
SOLLRIPOAVDMVAGAHVGLAGLAYSMVGNWAKVLLVMLLFAGVDGHRVSGGAA
ASDTRGLVLSFGSAQKIQLVNTGNSWHINRTALNCNDSLOTGFFAALFYKHKNSS
GCPERLASCRSIDKFAQWGPPLTYTPENSDORPCWYHAPRPGIIVPASOVCGPYVC
FPPSPVVGTTDRPGVTPYTWGANDSDVILLNTRPRGNWFCTWNGTGTCTCGG
FPCNIGAGNTLITCPTDCFRKHPKPEATYACGSGPWLTPRCWYHYPIRWHYPTVNF
TIFKVMYVGGVHREFAACNMTWRGCOLEDRSELSPLLLSTTEWQILPGLPSTTL
PALSTGLIHQNIQVQVLYGVGSVAVSLVWKVEYVLLFLLLADARICACLMMLL
IAQAEAALENLVNNAAVAGAGHTLSFLVFFCAAWVKGRLVPGAAYAFYGVWPLLL
LULALPPRAYA"
1. .2427

mat_peptide		434 a	745 c	714 g	540 t
BASE COUNT		434 a	745 c	714 g	540 t
ORIGIN		/product="unnamed"			
Query Match		96.5%	Score 619.8;	DB 6;	Length 2433;
Best Local Similarity		98.9%	Pred. No. 4.2e-145;		
Matches 624;		Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
QY	2	TGCCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTTACTGTCTGTCTGACCAATCCAG	61		
DB	506	TGCCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAG	565		
QY	62	CTTCCCGCTTATGAGTGCGCAACGTTGTCGGGATGACCATGTACAGAACGACTGTCTCCA	121		
DB	566	CTTCCCGCTTATGAGTGCGCAACGTTGTCGGGATGACCATGTACAGAACGACTGTCTCCA	625		
QY	122	ACTCAAGCATTGTGTATGAGGACGCGACATGATCATGACACACCCCGGGTGGCGCCT	181		
DB	626	ACTCAAGCATTGTGTATGAGGACGCGACATGATCATGACACACCCCGGGTGGCGCCT	685		
QY	182	GCCTTCGGGAGAACAACTCTTCCGCTGTGCGGTAGCGCTCACCCCGACGCTCGAGCTA	241		
DB	686	GCCTTCGGGAGAACAACTCTTCCGCTGTGCGGTAGCGCTCACCCCGACGCTCGAGCTA	745		
QY	242	GGAAACGCCAGCGTCCCGACACCAATACGACGCCACGTCGATTTGCTGTTGGGCGG	301		
DB	746	GGAAACGCCAGCGTCCCGACACCAATACGACGCCACGTCGATTTGCTGTTGGGCGG	805		
QY	302	CTGCTCTCTGTTCCGCTATGATGATGCGGGATCTCTGCGGATCTGCTCTGCTCTCC	361		
DB	806	CTGCTCTCTGTTCCGCTATGATGATGCGGGACCTCTGCGGATCTGCTCTGCTCTCC	865		
QY	362	AGCTGTTTACCAATCTCGCTTCGCGGATGACAGCGGTGACAGGACTGCAATTTGCTCAATCT	421		
DB	866	AGCTGTTTACCAATCTCGCTTCGCGGATGACAGCGGTGACAGGACTGCAATTTGCTCAATCT	925		
QY	422	ATCCCGGCCACATAACAGGTCACCGTATGGCTTTGGGATATGATGAACTGGTTCGCCTA	481		
DB	926	ATCCCGGCCACATAACAGGTCACCGTATGGCTTTGGGATATGATGAACTGGTTCGCCTA	985		
QY	482	CAACGGCCCTGGTGGTATCGACGTCGCTCGGATCCCAAGCTGTCGTGACATGGTGG	541		
DB	986	CAACGGCCCTGGTGGTATCGACGTCGCTCGGATCCCAAGCTGTCGTGACATGGTGG	1045		
QY	542	CGGGGGCCCATTTGGGAGTCTTGGCGGCTCTGCGGCTCTGCTACTATTCATGGTGGGAACCTGG	601		
DB	1046	CGGGGGCCCATTTGGGAGTCTTGGCGGCTCTGCGGCTCTGCTACTATTCATGGTGGGAACCTGG	1105		
QY	602	CTAAGGTTTTGATGTGATGCTACTCTTTTCG	632		
DB	1106	CTAAGGTTTTGATGTGATGCTACTCTTTTCG	1136		

RESULT 14
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM

AR157351
Sequence 49 from patent US 6245503.
AR157351
AR157351.1
GI:16218285

2433 bp
DNA
linear
PAT 17-OCT-2001

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2433)
TITLE	Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A. Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL	Patent: US 6245503-A 49 12-JUN-2001;
FEATURES	Location/Qualifiers
source	1. .2433
BASE COUNT	434 a 745 c 714 g 540 t
ORIGIN	/organism="unknown"
Query Match	96.5%; Score 619.8; DB 6; Length 2433;
Best Local Similarity	98.9%; Pred. No. 4.2e-145;
Matches 624;	Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	2 TGCCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTTACTGTCTGTGACCAATCCAG 61
Db	506 TGCCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAG 565
QY	62 CTTCCCGCTTATGAGTGCGCAACGTTGTCGGGATGTACCATGTACAGAACGACTGTCTCCA 121
Db	566 CTTCCCGCTTATGAGTGCGCAACGTTGTCGGGATGTACCATGTACAGAACGACTGTCTCCA 625
QY	122 ACTCAAGCATTGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGCGCCT 181
Db	626 ACTCAAGCATTGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGCGCCT 685
QY	182 GCCTTCGGGAGAACAACTCTTCCCGCTGTGCGGTAGCGCTCACCCCGACGCTCGAGCTA 241
Db	686 GCCTTCGGGAGAACAACTCTTCCCGCTGTGCGGTAGCGCTCACCCCGACGCTCGAGCTA 745
QY	242 GGAAACGCCAGCGTCCCGACACCAATACGACGCCACGTCGATTTGCTGTTGGGCGG 301
Db	746 GGAAACGCCAGCGTCCCGACACCAATACGACGCCACGTCGATTTGCTGTTGGGCGG 805
QY	302 CTGCTCTCTGTTCCGCTATGATGATGCGGGATCTCTGCGGATCTGCTCTGCTCTCCC 361
Db	806 CTGCTTCTGTTCCGCTATGATGATGCGGGACCTCTGCGGATCTGCTCTGCTCTCCC 865
QY	362 AGCTGTTTACCAATCTCGCTTCGCGGATGACAGCGGTGACAGGACTGCAATTTGCTCAATCT 421
Db	866 AGCTGTTTACCAATCTCGCTTCGCGGATGACAGCGGTGACAGGACTGCAATTTGCTCAATCT 925
QY	422 ATCCCGGCCACATAACAGGTCACCGTATGGCTTTGGGATATGATGAACTGGTTCGCCTA 481
Db	926 ATCCCGGCCACATAACGGGGTCACCGTATGGCTTTGGGATATGATGAACTGGTTCGCCTA 985
QY	482 CAACGGCCCTGGTGGTATCGACGTCGCTCGGATCCCAAGCTGTCTGTGACATGGTGG 541
Db	986 CAACGGCCCTGGTGGTATCGACGTCGCTCGGATCCCAAGCTGTCTGTGACATGGTGG 1045
QY	542 CGGGGGCCCATTTGGGAGTCTTGGCGGCTCTGCGGCTCTGCTACTATTCCATGGTGGGAACTGGG 601
Db	1046 CGGGGGCCCATTTGGGAGTCTTGGCGGCTCTGCGGCTCTGCTACTATTCCATGGTGGGAACTGGG 1105
QY	602 CTAAGGTTTTGATGTGATGCTACTCTTTTGC 632
Db	1106 CTAAGGTTTTGATGTGATGCTACTCTTTTGC 1136

RESULT 15
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX452798
Sequence 49 from Patent EP1211315.
AX452798
AX452798.1
GI:21712483

2433 bp
DNA
linear
PAT 06-JUL-2002

Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

PT proteins - in presence of di:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV

XX
ps Claim 23: Fig 21: 146pp; English.

AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

XX
sequence 642 BP. 109 A: 195 C: 176 G: 162 T; 0 other;

Query Match	100.0%	Score 642;	DB 17;	Length 642;
Best Local Similarity	100.0%	Pred. No. 1.3e-174;	Indels	0;
Matches 642;	Conservative	0;	Mismatches	0;

QY	1	ATGCCGGTTGCTCTTTTCTATCTTCTCTTGGCTTTACTGTCTGTGACCATTC	60
Db	1	ATGCCGGTTGCTCTTTTCTATCTTCTCTTGGCTTTACTGTCTGTGACCATTC	60
QY	61	GCTTCGGCTTATGAGGTGCGCAACGTGTCCGGATGTACCATGTACAGAACACTGTCTC	120
Db	61	GCTTCGGCTTATGAGGTGCGCAACGTGTCCGGATGTACCATGTACAGAACACTGTCTC	120
QY	121	AACTCAAGCATTTGTATGAGGACGCGACATGATCATGCAACGCCCGGTGGTGCCTC	180
Db	121	AACTCAAGCATTTGTATGAGGACGCGACATGATCATGCAACGCCCGGTGGTGCCTC	180
QY	181	TGGTTTCGGGAGAACACTCTTCCTCGCTCTGGGTAGCGCTCACCCCCACGCTGCGAGCT	240
Db	181	TGGTTTCGGGAGAACACTCTTCCTCGCTCTGGGTAGCGCTCACCCCCACGCTGCGAGCT	240
QY	241	AGGAACGCCAGCGTCCCCACCAAGCAATPACAGCGCAGTGCATTTGCTGTTGGGGCG	300
Db	241	AGGAACGCCAGCGTCCCCACCAAGCAATPACAGCGCAGTGCATTTGCTGTTGGGGCG	300
QY	301	GCTGCTCTCTGTTTCGGCTATGTACGTGGGGGATCTCGCGATCTGTCTTCCTGCTCTCC	360
Db	301	GCTGCTCTCTGTTTCGGCTATGTACGTGGGGGATCTCTGCGATCTGTCTTCCTGCTCTCC	360
QY	361	CAGTGTTTACCATCTCGCTCCCGGCGATGAGACGGTGCAGCATGCGAATTCCTCAATC	420
Db	361	CAGTGTTTACCATCTCGCTCCCGGCGATGAGACGGTGCAGCATGCGAATTCCTCAATC	420
QY	421	TATCCCGGCGACATACAGGTACCGGTATGGCTTGGGATATGATGATGAACCTGGTGCCT	480
Db	421	TATCCCGGCGACATACAGGTACCGGTATGGCTTGGGATATGATGATGAACCTGGTGCCT	480
QY	481	ACAAAGCGCCCTGGTGTATCGAGCTGCTCCGATCCCAACACTGTCGTGGACATGGTG	540
Db	481	ACAAAGCGCCCTGGTGTATCGAGCTGCTCCGATCCCAACACTGTCGTGGACATGGTG	540
QY	541	CGGGGGGCCATTCGGGAGTCTCGGCGGGCTCGCCTACTATTTCATGTGGGGAACTGG	600
Db	541	CGGGGGGCCATTCGGGAGTCTCGGCGGGCTCGCCTACTATTTCATGTGGGGAACTGG	600
QY	601	GCTAAGGTTTTGATTTGTGATGCTACTCTTTGCTCTCTTAATAG	642

601 GCTAAGGTTTGGATTGTGATGCTACTCTTTGCTCTCTAATAG 642

RESULT 2

AAI48913
ID AAI48913 standard: DNA; 642 BP.

AA
AC
XX
XX
DT
XX

XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW viricide; immunostimulant; vaccine; ds.

XX
OS

XX PN WO20025548-A2.

XX
18-JUL-2002.
PD

11-JAN-2002: 2002WO-EP00219.

XX
11-JAN-2001: 2001US-260699P.
PP

PR 30-AUG-2001; 2001US-313/68F.
yy

PA (INNO-) INNOGENETICS NV.

PI Maertens G, Bosman F, Buyse M,

DR WPI; 2002-599657/64.

XX

PT recombinant hepatitis

PT from HCV infection

XX Example 2; page 158-159; 243pp; English.
PS

xx CC The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1
CC and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.

XX
CC
commence 642 BP. 109 A: 195 C: 176 G: 162 T: 0 other;

100 0%. Score 642; DB 24; Length 642;

Best Local Similarity 100.0%; Pred. No. 1.3e-174;

Matches 64Z; Conservative 60

QY

Db

61 GCTTCGGCTTATGAGGTGCGCAACGTGTCGGGAGTACCAATCGACGAGT

Db 61 GCTTCCGCTTATGAGGTGCGCAACGTGTCCGGGATGTACCAATGTCACGAGCGGCTGCTG

QY
121 AACTCAAGCATTTGTGTATGAGGCAGCGACATGATCATTGACACCCCGGGTGGCTGG

Db
121 AACTCAAGCATGTGTATGAGGACGGACATGATCATGCACACCCCGGGGCGGCGGCCCC 18

181 TGGGTTGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCTCCACGCTCGCAGCT 24

SQ Sequence 2086 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

Query Match 96.5%; Score 619.8; DB 24; Length 2082;
Best Local Similarity 98.9%; Pred. No. 4.9e-168;
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGGCTTCTCTATCTTCCTCTTGCGCTTACTGTCTGCCTGACCATTCAG 61
DB |||||
QY 155 TGCCCGGTTGGCTTCTCTATCTTCCTCTTGCGCTTACTGTCTGCCTGACCATTCAG 214
DB |||||
QY 62 CTTCGCCCTTATGAGGTGCGCAACAGTGTCCGGGATGTACCATTGTACGAACAGCTGCTCCA 121
DB |||||
QY 215 CTTCGCCCTTATGAGGTGCGCAACAGTGTCCGGGATGTACCATTGTACGAACAGCTGCTCCA 274
DB |||||
QY 122 ACTCAAGCATTGTGTATGAGGCGGAGCATGATCATGCACACCCC CGGGTGCCT 181
DB |||||
QY 275 ACTCAAGCATTGTGTATGAGGCGGAGCATGATCATGCACACCCC CGGGTGCCT 334
DB |||||
QY 182 GCCTTCGGGAGAACAACTCTTCGGCTGTGGTAGCGCTACCCCCAAGCTCGCAGCTA 241
DB |||||
QY 335 GCCTTCGGGAGAACAACTCTTCGGCTGTGGTAGCGCTACCCCCAAGCTCGCAGCTA 394
DB |||||
QY 242 GGAAACGCCAGCTGCCACACGACAATAACGACGCGCATGCTGATTTGCTTGGGGCGG 301
DB |||||
QY 395 GGAAACGCCAGCTGCCACACGACAATAACGACGCGCATGCTGATTTGCTTGGGGCGG 454
DB |||||
QY 302 CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTCTTCCTCGTCTCCC 361
DB |||||
QY 455 CTGCTCTCTGTTCCGCTATGTACGTGGGGGAGCTCTGCGGATCTGTCTTCCTCGTCTCCC 514
DB |||||
QY 362 AGCTGTTTACCATCTCGCTTCGGCGCATGACACGGTGCAGGACTGCAATTGCTCAATCT 421
DB |||||
QY 515 AGCTGTTTACCATCTCGCTTCGGCGCATGACACGGTGCAGGACTGCAATTGCTCAATCT 574
DB |||||
QY 422 ATCCCGGCCACATAACAGGTCAACGCTACCGTATGGGATGATGATGAACCTGCGCTTA 481
DB |||||
QY 575 ATCCCGGCCACATAACAGGTCAACGCTACCGTATGGGATGATGATGAACCTGCGCTTA 634
DB |||||
QY 482 CAACGCCCTGTGGTATGCGAGCTGCTCCGGATCCCAAGCTGTCTGCGACATGCTGG 541
DB |||||
QY 635 CAACGCCCTGTGGTATGCGAGCTGCTCCGGATCCCAAGCTGTCTGCGACATGCTGG 594
DB |||||
QY 542 CGGGGCCCATTTGGGAGTCTTGGCGGCGCTCGCCTACTATTCCATGTTGGGGAACCTGG 601
DB |||||
QY 695 CGGGGCCCATTTGGGAGTCTTGGCGGCGCTCGCCTACTATTCCATGTTGGGGAACCTGG 754
DB |||||
QY 602 CTAAGGTTTTGATGTGATGCTACTCTTTGC 632
DB |||||
QY 755 CTAAGGTTTTGATGTGATGCTACTCTTTGC 785
DB |||||

RESULT 6
AAT12973
ID AAT12973 standard; DNA; 2086 BP.
XX AC
XX AAT12973;
DT XX
DT 24-SEP-1996 (first entry)
DE HCV E1 construct HCC165.
XX XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX OS Hepatitis C virus.
XX PN WO9604385-A2.
XX PD 15-FEB-1996.
XX PF 31-JUL-1995; 95WO-EP03031.
XX XX

PR Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;

Query Match 96.5%; Score 619.8; DB 17; Length 2086;
Best Local Similarity 98.9%; Pred. No. 4.9e-168;
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGGCTTCTCTATCTTCCTCTTGCGCTTACTGTCTGCCTGACCATTCAG 61
DB |||||
QY 155 TGCCCGGTTGGCTTCTCTATCTTCCTCTTGCGCTTACTGTCTGCCTGACCATTCAG 214
DB |||||
QY 62 CTTCGCCCTTATGAGGTGCGCAACAGTGTCCGGGATGTACCATTGTACGAACAGCTGCTCCA 121
DB |||||
QY 215 CTTCGCCCTTATGAGGTGCGCAACAGTGTCCGGGATGTACCATTGTACGAACAGCTGCTCCA 274
DB |||||
QY 122 ACTCAAGCATTGTGTATGAGGCGGAGCATGATCATGCACACCCC CGGGTGCCT 181
DB |||||
QY 275 ACTCAAGCATTGTGTATGAGGCGGAGCATGATCATGCACACCCC CGGGTGCCT 334
DB |||||
QY 182 GCCTTCGGGAGAACAACTCTTCGGCTGTGGTAGCGCTACCCCCAAGCTCGCAGCTA 241
DB |||||
QY 335 GCCTTCGGGAGAACAACTCTTCGGCTGTGGTAGCGCTACCCCCAAGCTCGCAGCTA 394
DB |||||
QY 242 GGAAACGCCAGCTGCCACACGACAATAACGACGCGCATGCTGATTTGCTTGGGGCGG 301
DB |||||
QY 395 GGAAACGCCAGCTGCCACACGACAATAACGACGCGCATGCTGATTTGCTTGGGGCGG 454
DB |||||
QY 302 CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTCTTCCTCGTCTCCC 361
DB |||||
QY 455 CTGCTCTCTGTTCCGCTATGTACGTGGGGGAGCTCTGCGGATCTGTCTTCCTCGTCTCCC 514
DB |||||
QY 362 AGCTGTTTACCATCTCGCTTCGGCGCATGACACGGTGCAGGACTGCAATTGCTCAATCT 421
DB |||||
QY 515 AGCTGTTTACCATCTCGCTTCGGCGCATGACACGGTGCAGGACTGCAATTGCTCAATCT 574
DB |||||
QY 422 ATCCCGGCCACATAACAGGTCAACGCTACCGTATGGGATGATGATGAACCTGCGCTTA 481
DB |||||
QY 575 ATCCCGGCCACATAACAGGTCAACGCTACCGTATGGGATGATGATGAACCTGCGCTTA 634
DB |||||
QY 482 CAACGCCCTGTGGTATGCGAGCTGCTCCGGATCCCAAGCTGTCTGCGACATGCTGG 541
DB |||||
QY 635 CAACGCCCTGTGGTATGCGAGCTGCTCCGGATCCCAAGCTGTCTGCGACATGCTGG 594
DB |||||
QY 542 CGGGGCCCATTTGGGAGTCTTGGCGGCGCTCGCCTACTATTCCATGTTGGGGAACCTGG 601
DB |||||
QY 695 CGGGGCCCATTTGGGAGTCTTGGCGGCGCTCGCCTACTATTCCATGTTGGGGAACCTGG 754
DB |||||
QY 602 CTAAGGTTTTGATGTGATGCTACTCTTTGC 632
DB |||||
QY 755 CTAAGGTTTTGATGTGATGCTACTCTTTGC 785
DB |||||

PA (INNO-) INNOGENETICS NV.

PI Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WFI; 1996-129401/13.

PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of di-sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

PS Claim 23; Fig 21; 146pp; English.

CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

XX

DEPT 5

RESULTS

AAT12973

ID AAT:

XX

AC AAT:

XX

24-1

22 23

XY

DE HCV

XX

KTW HCV

KW. Bere

KW 88

yy
yy
yy

25

OS : Hep:

PN WO91

XX

PD 15-

YY

22

31-0
PR

XX

635	CAACGGCCCTGTTGTTATCGAGCTGCTCCGGATCCCAAGCTTCGTGGACATGTTGG	694
Db		
542	CGGGGGCCATTGGGGAGTCTCTGGCGGGCCCTCGCCTACTATTCATGTTGGGGAATCGG	601
Qy		
695	CGGGGGCCCATTTGGGAGTCTCTGGCGGGCCCTCGCCTACTATTCATGTTGGGGAATCGG	754
Db		
602	CTAAGGTTTTGATTGTGATGCTACTCTTTGC	632
Qy		
755	CTAAGGTTTTGTTGTGATGCTACTCTTTGC	785
Db		

RESULT 7
AAT12974
IN AAT12974 standard; DNA: 2433 BP.

XX	
AC	AAT12974;
XX	
DT	25-SEP-1996 (first entry)

DE HCV E1 construct HCCI66.

XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.

XX
05
Hepatitis C virus.

XX
PN WO9604385-A2.

XX
PD
15-FEB-1996.XX
21 - III - 1995.

31-00F-1953, 53MO 210202Z
XX
0170 0000133

PR 29-JUL-1994; 94EP-08/013Z.
XX

PA (INNO-) INNOGENETICS NV.

XX
PT
Rosman F. Buyse M. De Martynoff G, Maertens G;

XX
XX
XX
DOI: 10.1002/1096-129401/13

Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - In presence of disulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

XX
PS Claim 23: Fig 21: 146pp; English.

XX CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV for *in vitro* detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2 and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

XX
SQ
Sequence 2433 BP: 434 A; 745 C; 714 G; 540 T; 0 other;

Query Match 96.5%; Score 619.8; DB 17; Length 2433;
Best Local Similarity 98.9%; Pred. No. 5.2e-168;
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	2	TGCCCGGTTGCTTTTCTCTATCTTCTCTTTGGGTTTACTGTCTCTGTCTGACCAATCCAG	61
Db	506	TGCCCGGTTGCTCTTTCTCTATCTTCTCTTTGGGTTTGTCTCTGTCTGACCGGTTTCAG	565
QY	62	CTTTCGGCTTATCAGGTGGCGCAACGTGTCGGGATGTACATGTCTACGAACGACTGCTCCA	121
Db	566	CTTTCGGCTTATAGTGTGGCAACGTGTCCGGATGTACCATGTACGGAACGACTGCTCCA	625
QY	122	ACTCAAGCAATTGTATGAGCAGCGGACATGATCATGCACACCCCCGGGTGCGTCCCT	181
Db	626	ACTCAAGCAATTGTATGAGCAGCGGACATGATCATGCACACCCCCGGGTGCGTCCCT	685
QY	182	GCCTTCGGGAGAACAACTCTTCCCGTGTCTGGGTAGCGTCAACCCCAAGCTCGCAGCTA	241
Db	686	GCCTTCGGGAGAACAACTCTTCCCGTGTCTGGGTAGCGTCAACCCCAAGCTCGCAGCTA	745
QY	242	GSAAACGACGCTGCCACACGACAATACGACGCCACGTCAATTTGCTTGTGTGGGGCGG	301
Db	746	GSAAACGACGCTGCCACACGACAATACGACGCCACGTCAATTTGCTTGTGTGGGGCGG	805
QY	302	CTGCTCTCTTTCCGCTATGTAGTGGGGGATCTCTGGGATCTGTCTTCCTCCTCTCCC	361
Db	806	CTGCTTCTCTGCTCCGCTATGTAGTGGGGGACCTCTGCGGATCTGTCTTCTCTGCTCCC	865
QY	362	AGCTGTTACCAATCTCGCCTCGCGGCATGAGACGGTGCAGGACTGCAATTTGCTCAATCT	421
Db	866	AGCTGTTACCAATCTCGCCTCGCGGCATGAGACGGTGCAGGACTGCAATTTGCTCAATCT	925
QY	422	ATCCCGGCCCAATAACAGGTCACCGTATGGCTTTGGGATATGATGAACTGTCGCCTTA	481
Db	926	ATCCCGGCCCAATAACGGGTACCGTATGGCTTTGGGATATGATGAACTGTCGCCTTA	985
QY	482	CAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCCAAGCTGTCTGTGGACATGGTGG	541
Db	986	CAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCCAAGCTGTCTGTGGACATGGTGG	1045
QY	542	CGGGGGCCCATTTGGGAGTCTCGGGGCTCCGCTACTATTCCATGGTGGGGAACTGGG	601
Db	1046	CGGGGGCCCATTTGGGAGTCTCGGGGCTCTGGCGGCTCTCGCTACTATTTCCATGGTGGGAACTGGG	1105
QY	602	CTAAGGTTTTGATTTGTATGCTACTCTTTTC	632
Db	1106	CTAAGTTTTTGGTTGTGAIGTACTCTTTTTC	1136

RESULT 8

AAI48940
ID AAI48940 standard: DNA: 2434 BP.

XX
ID
REF ID: A66400

AC AAL48940;

XX
DT 24-OCT-20XX
restitution arising from F2 prote

DE
heraillus c
vires p
XX

KW Hepatitis C virus; HCV; E1 protein; E2 protein; inbrection; gene
KW virocid: immunostimulant; vaccine; ds.
KW

XX
XX
XX

XX
OS
Нератитів с вітн

PN
XX
WO200255548-

PD 18-JUL-2002.
XX

11-JAN-2002; 2002WO-EP00219.

11-JAN-2001; 2001US-260699P.
PR
AA

30-AUG-2001; Z00105=313766F
PR XX

PA (INNO-) INNOGENETICS NV.
XX

PI Maertens G, Bosman F, Buyse M;

[illegible]

RESULT 10	
AAT30386	
ID	AAT30386 standard; cDNA; 3461 BP.
XX	
AC	AAT30386;
XX	
DT	22-AUG-1996 (first entry)
XX	
DE	5'UTR/CORE/ENV/NS1/NS2/NS3 cDNA from HCV (#4).
XX	
KW	Hepatitis C virus; HCV; antigen; detection; antibody; ds.
XX	
OS	Hepatitis c virus.
XX	
PH	Key
FT	Location/Qualifiers
FT	307..3461
FT	/*tag= a
FT	307..879
FT	/*tag= b
FT	/product= Core peptide
FT	880..1455
FT	/*tag= c
FT	/product= ENV1
FT	1456..2736
FT	/*tag= d
FT	/product= NS1/ENV2
FT	2737..3461
FT	/*tag= e
FT	/product= NS2 and NS3
XX	
XX	
PN	JP07133291-A.
XX	
PD	23-MAY-1995.
XX	
PF	18-JUN-1993;
XX	
XX	93JP-0147944.
XX	
PR	18-JUN-1993;
XX	
XX	93JP-0147944.

XX	(TOFU) TONEN CORP.
PA	
XX	WPI; 1995-220780/29.
DR	P-PSDB; AAR98361.
XX	
PT	Recombinant polypeptide comprising partial NS1 region of hepatitis
PT	non-A non-B viral antigen - used in a method for detecting
PT	antibodies against hepatitis non-A non-B virus.
XX	
XX	Disclosure; Page 10-12; 15pp; Japanese.
XX	
CC	The sequences given in AAT30386-87 encode the 5'UTR/CORE/ENV/NS1/NS2/ NS3 protein region derived from hepatitis C virus (HCV) isolates #4 and #6 respectively. The proteins encoded by these sequences partic. encode amino acids 384-495 of the HCV NS1 antigen. These protein fragments may be used in the detection of antibodies against HCV.
XX	
SQ	Sequence 3461 BP; 638 A; 1046 C; 1012 G; 765 T; 0 other;
	Query Match 88.2%; Score 566.4; DB 16; Length 3461;
	Best Local Similarity 93.5%; Pred. No. 1.4e-152;
	Matches 591; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY	2 TGCCGGTTGCTTTTCTCTATCCTTCCTGTGCCTTAAGTGCTCTGACATTCCAG 61
Db	812 TGCCGGTTGCTTTTCTCTATCCTTCCTGTGCTTCTGCTGTGACCATCCCAG 871
QY	62 CTTCCTCGTTATGAGGTGCGCAAGCTGTCCGGGATGACATGTACGAACGACTGCTCCA 121
Db	872 CTTCCTCGTTATGAGGTGCGCAAGCTGTCCGGGATGACATGTACGAACGACTGCTCCA 931
QY	122 ACTCAAGCATGTGTATGAGGACGGACATGATCATCACACCCCGGGTGGTGCCTCCT 181
Db	932 ACTCAAGTATGTGTATGAGGACGGACGATGATCATGACACCCCGGGTGGTGCCTCCT 991
QY	182 GCGTTCCGGAGAAACAATCTTTCCCGTGTGTGGGTAGCGCTACGCCCAAGCTCGCAGCTA 241
Db	992 GCGTTCCGGAGAGCAATTTCTCCGCTGTGTGGGTAGCGCTACGCCCAAGCTCGCAGCTA 1051
QY	242 GGAAGCCAGCGTCCCCACACAGCAATACGACCGCCAGCTCGATTTGCTGTGGGGCGG 301
Db	1052 GAAACAGCAGCATCCCCACTACGACAATACGCCCATGTCGATTTGCTGTGGGGCGAG 1111
QY	302 CTGCTCTCTGTTCCGCTATGATGATGGGGGATCTCTCGGATCTGTCTTCTCGTCTCCC 361
Db	1112 CTGCTCTCTGCTCCGCATGTACGTGGGGATCTCTCGGATCTGTCTTCTCGTCTCCC 1171
QY	362 AGCTGTTACCATCTCGCTCGCGGATGAGAGCGGTGAGGACTGCAATTGCTCAATCT 421
Db	1172 AGCTGTTACCATCTCTACCTCGCGGATGAGAGCGGTGAGGACTGCAATGCTCAATCT 1231
QY	422 ATCCCGGGCCATAACAGGTACCGTATGGCTTCGGGATGATGATGAATGAACTGTCGCTA 481
Db	1232 ATCCCGGGCCAGTGTGAGTGTGCGGATGCGGATGCGGATGATGATGAATGAACTGTCGCTA 1291
QY	482 GAAAGGCCCTGTGTATGACAGCTGTCTCCGATGCCAACAGCTGTGCTGACATGCTGG 541
Db	1292 CAACAGCCCTGTGTATGACAGCTGTCTCCGATGCCAACAGCTGTGACATGCTGG 1351
QY	542 CGGGGGCCCATTTGGGAGTCTTGGGGGCTCGGCTACTATTCAATGTTGGGAACTGGG 601
Db	1352 CAGGGGCCCATTTGGGAGTCTTGGGGGCTCGGCTACTATTCAATGTTGGGAACTGGG 1411
QY	602 CTAAGGTTTGATGTGATGCTACTCTTTGCT 633
Db	1412 CTAAGGTTTGATGTGATGCTACTCTTTGCT 1443

RESULT 11
AAQ24467
ID AAQ24467 standard; DNA; 1880 BP.
XX

Mon Dec 22 13:28:48 2003

us-09-899-303a-3.rng

Best Local Similarity 93.5%; Pred. No. 2.2e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTTCTCTAATCTTCTCTTGGCTTTACTGTCCTGTCTGACCATTCACG 61
Db 506 TGCCCGGTGCTCTTTCTCTAATCTTCTCTTGGCTTTACTGTCCTGTCTGACCATTCACG 565

QY 62 CTTCCCGCTTATGAGTGCAGCAAGTGTCCGGGATGTACCATGTTCACGAACGACTGCTCCA 121
Db 566 CTTCCCGCTTATGAGTGCAGCAAGTGTCCGGGATGTACCATGTTCACGAACGACTGCTCCA 625

QY 122 ACTCAAGCATTTGTATGAGGACGAGACATGATATGACACACCCCGGCTGGCTTCCCT 181
Db 626 ACTCAAGCATTTGTATGAGGACGAGACATGATATGACACACCCCGGCTGGCTTCCCT 685

QY 182 CGCTTCGGGAGAACATCTTCCCGCTGCTGGGTAGCGCTCACCCGACGCTCGCAGCTA 241
Db 686 CGCTTCGGGAGAACATCTTCCCGCTGCTGGGTAGCGCTCACCCGACGCTCGCAGCTA 745

QY 242 GGAACGCGCAGGTCCCGACCAACGACATACGACGCCACGTCGATTTGCTGTTGGGGCGG 301
Db 746 GGAATGCCAGGTCCCGACCAACGACATACGACGCCACGTCGATTTGCTGTTGGGGCGG 805

QY 302 CTGCTCTCTGCTTCGCTATGATGATGAGGAGATCTCTGGGATCTGCTCTCTCTCTCC 361
Db 806 CTGCTCTCTGCTTCGCTATGATGATGAGGAGATCTCTGGGATCTGCTCTCTCTCTCC 865

QY 362 AGCTGTTCCACCATCTCGGCTCGCGGATGAGAGCGGTGCAGGACTGCAATTTGCTCAATCT 421
Db 866 AGCTGTTCCACCATCTCGGCTCGCGGATGAGAGCGGTGCAGGACTGCAATTTGCTCAATCT 925

QY 422 ATCCCGGCGCACATAACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 481
Db 926 ATCCCGGCGCACATAACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 985

QY 482 CAAAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
Db 986 CAAAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1045

QY 542 CGGGGGCCCATTTGGGAGTCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 601
Db 1046 CGGGGGCCCATTTGGGAGTCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 1105

QY 602 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632
Db 1106 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136

RESULT 13
AAQ43889
ID AAQ43889 standard; cDNA to mRNA; 2540 BP.
XX
AC AAQ43889;
XX
DT 21-OCT-1993 (first entry)
XX
DE NANB hepatitis virus polynucleotide N-2540-2.
XX
KW Non-A, non-B; virus; polymerase chain reaction; detection;
KW sensitive; specific; HCV; NANBH; ss.
XX
OS Non-A, non-B hepatitis virus.
XX
FH Key Location/Qualifiers
FH CDS 342..2540
FT /*tag= a
FT 1..341
FT /*tag= b
FT /note= "from 5' terminal of NANBH virus RNA"
XX
PN JP05091884-A.
XX
PD 16-APR-1993.

XX 10-APR-1991; 91JP-0196175.
XX 12-JUN-1990; 90JP-0153401.
XX 08-NOV-1990; 90JP-0304405.
XX (NAKA/) NAKAMURA T.
XX WPI; 1993-199637/25.
XX P-PSDB; AAR38279.
XX Antigen related to non-A and non-B hepatitis virus - comprises
PT non-translation region comprising 340 - 341 moles. of nucleotides,
PT non-translation region comprising 1885 - 2551 moles. of
PT nucleotides including region 1,149 and, etc.
XX Claim 3; Page 19-20; 73pp; Japanese.
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
CC which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC The polypeptide it encodes may be used in a system for detecting
CC NANB hepatitis. This method is highly specific and sensitive, and
CC can detect NANB hepatitis virus which could not be detected by
CC conventional methods.
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
SQ

Query Match 88.1%; Score 565.4; DB 14; Length 2540;
Best Local Similarity 93.5%; Pred. No. 2.4e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTTCTCTAATCTTCTCTTGGCTTTACTGTCCTGTCTGACCATTCACG 61
Db 847 TGCCCGGTGCTCTTTCTCTAATCTTCTCTTGGCTTTACTGTCCTGTCTGACCATTCACG 906

QY 62 CTTCCCGCTTATGAGTGCAGCAAGTGTCCGGGATGTACCATGTTCACGAACGACTGCTCCA 121
Db 907 CTTCCCGCTTATGAGTGCAGCAAGTGTCCGGGATGTACCATGTTCACGAACGACTGCTCCA 966

QY 122 ACTCAAGCATTTGTATGAGGACGAGACATGATGATGATGATGATGATGATGATGATGAT 181
Db 967 ACTCAAGCATTTGTATGAGGACGAGACATGATGATGATGATGATGATGATGATGATGAT 1026

QY 182 GCGTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCGCTCGCAGCTA 241
Db 1027 GCGTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCGCTCGCAGCTA 1086

QY 242 GGAACGCGCAGGTCCCGACCAACGACAAATACGACGCAACGCTCGACTTGTGTTGGGGCGG 301
Db 1087 GGAATGCCAGGTCCCGACCAACGACAAATACGACGCAACGCTCGACTTGTGTTGGGGCGG 1146

QY 302 CTGCTCTCTGCTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
Db 1147 CTGCTCTCTGCTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206

QY 362 AGCTGTTCCACCATCTCGCTTCGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
Db 1207 AGCTGTTCCACCATCTCGCTTCGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266

QY 422 ATCCCGGCGCACATAACAGGTACCGTATGCTTGGGATGATGATGATGATGATGATGATGAT 481
Db 1267 ATCCCGGCGCACATAACAGGTACCGTATGCTTGGGATGATGATGATGATGATGATGATGAT 1326

QY 482 CAAAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
Db 1327 CAAAGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386

QY 542 CGGGGGCCCATTTGGGAGTCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 601
Db 1387 CGGGGGCCCATTTGGGAGTCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCT 1446

QY 602 CTAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632

Db 1447 CTAAGTCTGATTGTGGCGCTACTCTTCGC 1477

AAQ63753
RESULT 14
ID AAQ63753 standard; cDNA to mRNA; 2540 BP.
XX
AC AAQ63753;
XX
DT 30-JAN-1995 (first entry)
XX
DE NANBHV genomic fragment #2.
XX
KW Polymerase chain reaction; PCR; primer; amplify; detection; NANBHV;
KW non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
XX
OS Synthetic.
XX
PN JP06125777-A.
XX
PD 10-MAY-1994.
XX
PF 20-JUN-1991; 91JP-0247120.
XX
PR 20-JUN-1991; 91JP-0247120.
XX
PA (NAKA/) NAKAMURA T.
XX
DR WPI; 1994-187937/23.
XX
PT Oligonucleotide primer pairs specific for non-A, non-B hepatitis
PT virus - used for high sensitivity detection of non-A non-B (NANB)
PT hepatitis virus
XX
PS Disclosure; Page 24-25; 25pp; Japanese.
XX
CC The sequences given in AAQ63752-53 represent fragments of the non-A,
CC non-B hepatitis virus (NANBHV) genome. These fragments were amplified
CC using the primers given in AAQ63732-51. These primers were used in the
CC detection of NANBH. The primers are based on the 5'-terminal region and
CC the core protein coding region. The method allows highly sensitive
CC detection of NANBHV.
XX
SQ Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

Query Match 88.1%; Score 565.4; DB 15; Length 2540;
Best Local Similarity 93.5%; Pred. No. 2.4e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTATCTTCTCTTCTGCTTTACTGTCTGTGACCAATCCAG 61
DB 847 TGCCCGGTTGCTCTTTCTATCTTCTCTTGGCTTTGCTGTGTTGACCATCCAG 906

QY 62 CTTCCGCTTATGAGTGCGCAAGTGTCCGGATGTACCATGTCAGAACACTGTCTCA 121
DB 907 CTTCCGCTTATGAGTGCGCAAGTGTCCGGATGTACCATGTCAGAACACTGTCTCA 966

QY 122 ACTCAAGCATTTGTATGAGGACGAGCATGATGACACCCCGGCTGCGCCT 181
DB 967 ACTCAAGCATTTGTATGAGGACGAGCATGATGACATCTCCCGGCTGCGCCT 1026

QY 182 GCGTTCCGGAGAACAACTCTTCCCGCTGTGCTGAGCGCTTCACCCCGCATCGCAGCTA 241
DB 1027 GCGTTCCGGAGAACAACTCTTCCCGCTGTGCTGAGCGCTTCACCCCGCATCGCAGCTA 1086

QY 242 GGAACGCGAGCTCCCAACACACAAATACGACGCACTGATTTGCTGTTGGGCGG 301
DB 1087 GGAATGCCAGCTGCCCACTACGACAAATACGACGCACTGATTTGCTGTTGGGCGG 1146

QY 302 CTGCTCTCTGTTCCGCTATGATGAGGAGATCTTCCGATCTGTTCTCTGTTCC 361
DB 1147 CTGCTCTCTGTTCCGCTATGATGAGGAGATCTTCCGATCTGTTCTCTGTTCC 1206

QY 362 AGCTGTTACCAATCTCGCTCGCGGATGAGCGGTGCGAGCTGCAATTTGCTCAATCT 421
DB 1207 AGCTGTTACCAATCTCGCTCGCGGATGAGCGGTGCGAGCTGCAATTTGCTCAATCT 1266

QY 422 ATCCCGGCGCACATAACAGGTCAACGCTTGGCTTGGGATATGATGATGAAGTGTGCGCTA 481
DB 1267 ATCCCGGCGCATTTATCAGGTCAACGCTTGGCTTGGGATATGATGATGAAGTGTGCGCTA 1326

QY 482 CAAAGGCTTGGTGTATCGAGCTCTCGGATCCCAAGCTGTGCGGACATGCTGG 541
DB 1327 CAAAGGCTTGGTGTATCGAGCTCTCGGATCCCAAGCTGTGCGGACATGCTGG 1386

QY 542 CGGGGCGCATTTGGGAGTCTCGCGGCTCGCTACTATTCCATGTTGGGAACTGGG 601
DB 1387 CGGGGCGCATTTGGGAGTCTCGCGGCTCGCTACTATTCCATGTTGGGAACTGGG 1446

QY 602 CTAAGGTTTGTATGATGATGCTACTCTTTGC 632
DB 1447 CTAAGGTTTGTATGATGATGCTACTCTTTGC 1477

RESULT 15
ABK91411
ID ABK91411 standard; DNA; 9605 BP.
XX
AC ABK91411;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
XX
PN WO200259321-A2.
XX
PD 01-AUG-2002.
XX
PF 16-JAN-2002; 2002WO-EP00526.
XX
PR 23-JAN-2001; 2001US-263479P.
XX
PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
PI De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
DR P-PSDB; ABG32451.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression -
XX
PS Claim 9; Page 36-39; 69pp; English.
XX
CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an

CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is the HCV replicon Con 1, used as a basis for
CC the adaptive mutations of the invention.

XX
XX
SQ Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 other;

Query Match 88.1%; Score 565.4; DB 24; Length 9605;
Best Local Similarity 93.5%; Pred. No. 3.9e-152;
Matches 590; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTTCTCTGCTTTACTGTCTGTCTGACCATTCACG 61
DB |||||
QY 847 TGCCCGGTGCTCTTCTCTATCTTCTCTGCTTTACTGTCTGTCTGACCATTCACG 906
DB |||||
QY 62 CTTCCTGCTTATGAGTGGCGCAAGTGTCCGGGATGTACCATGTACGAGAGCTGTCCA 121
DB |||||
QY 907 CTTCCTGCTTATGAGTGGCGCAAGTGTATCCGAGGTGTACCATGTACGAGAGCTGTCCA 966
DB |||||
QY 122 ACTCAAGCATGTGTATGAGGACGCGGACATGATCATGCACACCCCGGTGGTGCCT 181
DB |||||
QY 967 ACCTAAGCATGTGTATGAGGACGCGGACATGATCATGCATACCCCGGTGGTGCCT 1026
DB |||||
QY 182 GCGTTCGGGGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTA 241
DB |||||
QY 1027 GCGTTCGGGGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTA 1086
DB |||||
QY 242 GGAACGCGCAGCGTCCCGACACGACATACGACCCACGCTCGATTTGCTCGTTGGGCGG 301
DB |||||
QY 1087 GGAACGCTAGCGTCCCGACACGATACGACCCCATGTGATTTGCTCGTTGGGCGG 1146
DB |||||
QY 302 CTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTCTTCTCTCTCC 361
DB |||||
QY 1147 CTGCTCTCTGTTCCGCTATGTACGTGGGATCTCTCGGATCTGTCTTCTCTCTCC 1206
DB |||||
QY 362 AGCTGTTTCCATCTCGCTCCCGGCATGAGACGCTGAGGACTGCAATTTGCTCAATCT 421
DB |||||
QY 1207 AGCTGTTTCCATCTCTCGCTCCCGGCACGAGACAGTACAGGACTGCAATTTGCTCAATAT 1266
DB |||||
QY 422 ATCCCGGCACATACAGGTACCGTATGCTGGGATGATGATGAATGCTGGTCCGCTA 481
DB |||||
QY 1267 ATCCCGGCACATGAGGTACCGTATGCTGGGATGATGATGAATGCTGGTCCGCTA 1326
DB |||||
QY 482 CAACGGCCCTGCTGATCGAGCTGCTCCGGATCCCAAGCTGTGCTGGACATGGTGG 541
DB |||||
QY 1327 CAGGAGCCCTAGTGGTATCGAGTTACTCCGGATCCCAAGCTGTGCTGGATATGGTGG 1386
DB |||||
QY 542 CGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCTACTATTCCATGCTGGGGAACCTGGG 601
DB |||||
QY 1387 CGGGGGCCCATTTGGGAGTCTTGGGGGCTTGGGCTTCTATTCCATGCTGGGGAACCTGGG 1446
DB |||||
QY 602 CTAAGGTTTGTGATGTGCTACTCTTTGC 632
DB |||||
QY 1447 CTAAGGTTTGTGATGTGCTACTCTTTGC 1477
DB |||||

Search completed: December 19, 2003, 18:51:00
Job time : 182.486 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1651.58 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCCGTGTCTTCTC.....TACTCTTGTCTCTCTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_esti:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	117.2	18.3	488	9	AV755731
C 2	92.6	14.4	492	9	AV758366
C 3	41.6	6.5	534	14	CD040840
C 4	40.6	6.3	925	29	CNS0091P
					AL053013 Drosophil

5	40.4	6.3	590	14	CA659369
6	40.4	6.3	645	29	CNS01213
7	40	6.2	525	10	BE337089
8	40	6.2	671	12	BI723733
9	40	6.2	740	12	BJ536071
10	40	6.2	1201	13	EX356664
11	39.6	6.2	423	9	AA459034
12	39	6.1	621	14	CA816001
13	39	6.1	624	12	BI723734
14	38.6	6.0	399	9	AV638521
15	38.6	6.0	434	9	AV637507
16	38.6	6.0	440	9	AV637983
17	38.6	6.0	450	9	AV637259
18	38.6	6.0	451	9	AV637328
19	38.6	6.0	451	9	AV637643
20	38.6	6.0	453	9	AV634724
21	38.6	6.0	454	9	AV637050
22	38.6	6.0	456	9	AV635382
23	38.6	6.0	473	9	AV632765
24	38.6	6.0	481	9	AV635503
25	38.6	6.0	485	9	AV632811
26	38.6	6.0	506	9	AV392445
27	38.6	6.0	508	9	AV634095
28	38.6	6.0	526	9	AV641895
29	38.6	6.0	533	9	AV638125
30	38.6	6.0	537	9	AV632335
31	38.6	6.0	588	9	AV387329
32	38.2	6.0	431	9	AV639153
33	38.2	6.0	501	9	AV638474
C 34	38	5.9	742	13	BQ752673
35	38	5.9	1195	28	BI0902
36	37.8	5.9	435	14	C72860
C 37	37.8	5.9	634	10	BF396591
38	37.4	5.8	497	9	AV633658
39	37.4	5.8	856	29	BZ578381
40	37.4	5.8	872	29	BZ55011
41	37.2	5.8	888	10	BG365821
42	37.2	5.8	1201	13	BX420106
43	37	5.8	610	14	CB657655
C 44	37	5.8	945	29	CNS05KMQ
45	36.8	5.7	431	9	AV636681

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 BM Homo sapiens CDNA clone BMFAK03 5', mRNA linear EST 19-OCT-2000
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens CDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangliang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

[illegible]

```
/tissue type="infected host tissue"
/cell_line="P6497"
/dev_stage="48 hour post infection"
/clone_lib="psHB: Infected hypocotyl soybean host. 48 hrs
post infection"
/note="Vector: pBK-CMV, Site 1: EcoRI; Site 2: XhoI;
USDA-IFAPS: Expression of Phytophthora sojae genes during
infection and propagation."
BASE COUNT 101 a 187 c 159 g 87 t
ORIGIN

Query Match 6.5%; Score 41.6; DB 14; Length 534;
Best Local Similarity 47.1%; Pred. No. 2.5;
Matches 128; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 70 TATGAGTGCACACGTGTCGGGATGTATACCATGTACGAAACGACTGCTCCAACTCAAGC 129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 TAGCGGTGCGGAGATTTACGGTATCCGATCGGCTTCGCCGGCTTCTACAACTGGACC 259
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 130 ATTGTATGAGCAGCAGCATATGACACACCCCGGGTGTGCGCTCGCTCGTTCGG 189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 TCGATGGACCAAGAGAGGCCCGATCATGCTGACCCCAAGACGGTGGCCAACTTAC 319
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 190 GAGAACTCTTCCCGCTGCTGGGTAGCGCTACCCCGACGCTCGCAGTAGGACGCC 249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 CACTAGCGCGGACCACTCTCGGCTCGAACGGTGGCTTCGACGTGACAAAGATTATC 379
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 250 AGCGTCCCAACGACCAATATGACGCCACGCTCGATTGCTGCTGGGGCGGCTGCTCTC 309
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 AACTTCTGACGACGAACGGCTCTCGCAGGTATAGTGTGCGGCTGACGCGACCCAC 439
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 310 TGTTCGGCTATGATCGTGGGGATCTCTCGG 341
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 CGTCCGCAACAAGATCTCGGAGGAGTCCG 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013
VERSION 1 (bases 1 to 925)
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1. .925
FEATURES
source

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR19D16"
/clone_lib="RPI-98"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 6.3%; Score 40.6; DB 29; Length 925;
Best Local Similarity 15.8%; Pred. No. 5.5;
Matches 69; Conservative 165; Mismatches 198; Indels 4; Gaps 1;

QY 122 ACTCAAGCATTTGTATGAGGAGCGGACATGATGACACCCCGGGTGTGCGCTT 181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 AATANANNNTTATTTAANNNNANANANANANANANANANANANANANANANANAN 548
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 182 GCGTTCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCGACGTA 241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 549 TTTTSSSGYGKCGSGSGHSCSCSCSSSCSCBCCCCSCSYCCSSBSBS 608
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 242 GGAACCCAGCGTCCCAACGACAAATACGACGCGCATGATGCTGCTGGGGCGG 301
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 609 KCSSTBSBSCCCSKSVGTSCSS----SSSCSSSSSTSSSTSSSTSSSTSSSS 664
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 302 CTGCTCTGTTCCGCTATGATGCTGGGGATCTCTCGGATCTGCTTCTCTCTCC 361
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 SSSSYTTSKTSAGSGSWSAGGGSGTGTSSSSSSTSTSSSVSSSGKSTBSGS 724
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 362 AGCTGTTACCATCTCGCTCGCGCATGAGCGGTGAGGACTGCAATTGCTCATCT 421
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 BSSSGSSSSSTSSBBSCTSTSSSSSSSSTSCCTCCCSYSSSTSSSTSSSTSS 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 422 ATCCCGGCACATAACAGGTCAACGCTGCTGGGATATGATGATGAACGCTGCGCTA 481
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 SGSSSSSVGTSSSDSTSTCCSCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 844
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 482 CAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CGCGSSSTNGMBGTSSACSSSSSSSCSSSSSSSSSSSSSSSSSSSSSSSSSS 904
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 542 CGGGGCGCCATGGGG 557
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 KSSSGSVSSGSGSGS 920
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
CNS0091P 590 bp mRNA linear EST 24-NOV-2002
LOCUS wlm1.pk0005.c9 wlm1 Triticum aestivum cDNA clone wlm1.pk0005.c9 5'
DEFINITION end, mRNA sequence.
ACCESSION CA659369
VERSION 1 GI:25237894
KEYWORDS Triticum aestivum (bread wheat)
SOURCE Triticum aestivum
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 590)
Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
Miao, G., Caraher, N. and Hanafey, M.K.
DuPont Wheat cDNA Sequence
Unpublished
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
Location/Qualifiers
1. .925
FEATURES
source
```

```
source
1. .590
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Stephens"
/db_xref="taxon:4565"
/clone="wlm1.pk0005.c9"
/tissue_type="leaf"
/notes="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
XhoI; Wheat (Triticum aestivum L.) seedlings 1 hr after
inoculation with Erysiphe graminis f. sp tritici"
BASE COUNT      106 a      218 c      159 g      88 t      19 others
ORIGIN

Query Match      6.3%; Score 40.4; DB 14; Length 590;
Best Local Similarity 47.3%; Pred. No. 5.3;
Matches 122; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 147 GGACATGATCATCACACCCCGGTGGTGGTTCGGGAGAACAACTCTTCGCG 206
Db 43 GTAGATGAGCATGGCGGCGACGCGGCCACGGCTGGCCCTCCACCTCTCTTC 102
QY 207 CTGCTGGGTAGCGCTACCCGCCACGCTCGCAGCTAGGAAAGCGCAGCGTCCCCACACAGAC 266
Db 103 CGTCTCGCCCTCGACCGCCGACACAGCCACCCGAGCCCTGCGCCTCGCTCCG 162
QY 267 AATACAGACGCAGCTGATTGCTGTTGGGGGGCTGCTCTCTGTTTCGGTATGTACT 326
Db 163 CTTCCTTGGCGCGGCTCACCTCGCGGCCGCCACAGACGCGCCCTCCCGCTCTCCGT 222
QY 327 GGGGATCTCGGGATCTGCTCTCTGCTCCAGCTGTTCCACATCTCGGCTCGCG 386
Db 223 CTCGCTGGCGGCGCTGGAGCGGCCCTCTCTCTGCGCGCTCATATGACGAGCGCG 282
QY 387 GCATGAGACGTTGCAGGA 404
Db 283 CAAGTCAAGTGAAGCA 300

RESULT 6
CNS01213      645 bp      DNA      linear      GSS 26-JUL-1999
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence 17 end of BAC
BACN08C07 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL101589
VERSION
AL101589.1      GI:5613200
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 645)
REFERENCE
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
Location/Qualifiers
source
1. .645
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN08C07"

/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/notes="end : T7"
BASE COUNT      28 a      26 c      85 g      92 t      414 others
ORIGIN

Query Match      6.3%; Score 40.4; DB 29; Length 645;
Best Local Similarity 10.9%; Pred. No. 5.5;
Matches 28; Conservative 85; Mismatches 144; Indels 0; Gaps 0;

QY 351 CCTGCTCCAGCTGTTTCACATCTCCCTCGCGGCATGAGCGGTGACAGTGTGCA 410
Db 191 SCCSCNCTCNNTNTTNTSSSSSSNNNTNNNNNNNNNNNNNNNNNNNNNNNNNS 250
QY 411 TTGCTCAATCTATCCCGGCCACATAACAGGTACCGTATGCTTGGGATATGATGAA 470
Db 251 GGSSNNNSNNNGSSSSSSNTSSNNNNNNNNNNNNNNNNNNNNNNNNNNSSSTS 310
QY 471 CTGGTTCGCTACAAAGCCCTGTTGTTATCGAGCTGCTCGGATCCCAAGTGTGCT 530
Db 311 TTNTSNTSSNNATGSSSSSSSGTTTBSGSSSSSSSSNNNNNNNNNNNNNNNN 370
QY 531 GGACATGTTGGCGGGGCCATTTGGGAGTCTGCGGGCTCGCTACTTCCATGCT 590
Db 371 SNSATTSSTNGSSSSSSSSSTSTSTSTSSNNNTSTSTSSSTTNTSTSTSS 430
QY 591 GGGGAACCTGGGCTAAGG 607
Db 431 SSNSTGSSSTTTTNS 447

RESULT 7
CNS01213      525 bp      mRNA      linear      EST 14-JUL-2000
LOCUS
DEFINITION
Chlamydomonas reinhardtii CC-1690, normalized, Lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION
BE337089
VERSION
BE337089.1      GI:9210174
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 525)
REFERENCE
Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
Unpublished
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlamy@duke.edu.
Location/Qualifiers
source
1. .525
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
```


ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 99 a 181 c 160 g 85 t
ORIGIN

Query Match 6.2%; Score 40; DB 10; Length 525;
Best Local Similarity 47.0%; Pred. No. 6.5;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGGAGGACATCATGACACACCCC 167
Db 43 GCACCCCTTCAACAACCCACCGTGTGAATTCGCCCGCTCAAGTACCTGTGTCAT 102
QY 168 CGGGTGGTGGCTGCTGGGAGAACAACTTCCCGCTGTGGGTAGCGCTCACCCC 227
Db 103 GGAGGACAAGACCTGCACCTGGAGTCTCGAGTACTGGACCTGGGCAACCTGTCCAA 162
QY 228 CAGCTCGGAGTAGGAAGCGCAGCTGCCACACAGCAATACGACGCGCACGTCGATTT 287
Db 163 CGCGCTCAAAAACAACATCTTCATGATCCCAACCCCGTTCATCGCGCGCGCGCGCGC 222
QY 288 GCTCGTTGGGCGGCTGCTCTGTTCGGTATGTAGTGGGGATCTCTGCGGATCTGT 347
Db 223 GGGCGAGCGCGCGCGCGCGAGAGTAGCGGAGCGCGCGCGAGCCCAATGAAGT 282
QY 348 CTTCTCTGCTCCAGCTGTTCAC 371
Db 283 CAACATGCGCACCTGCTCTCAC 306

RESULT 8
BI723733 671 bp mRNA linear EST 19-SEP-2001
LOCUS 1031067P08.y1 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI723733
VERSION BI723733.1 GI:15699428
KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 671)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031

JOURNAL Unpublished
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source
1. .671
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant

Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

BASE COUNT 139 a 224 c 209 g 99 t
ORIGIN

Query Match 6.2%; Score 40; DB 12; Length 671;
Best Local Similarity 47.0%; Pred. No. 7.1;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 108 GAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGGAGGACATCATGACACACCCC 167
Db 353 GCACCCCTTCAACAACCCACCGTGTGAATTCGCCCGCTCAAGTACCTGTGTCAT 412
QY 168 CGGGTGGTGGCTGCTGGGAGAACAACTTCCCGCTGTGGGTAGCGCTCACCCC 227
Db 413 GGAGGACAAGACCTGCACCTGGAGTCTCGAGTACTGGACCTGGGCAACCTGTCCAA 472
QY 228 CAGCTCGGAGTAGGAAGCGCAGCTGCCACACAGCAATACGACGCGCACGTCGATTT 287
Db 473 CGCGCTCAAAAACAACATCTTCATGATCCCAACCCCGTTCATCGCGCGCGCGCGCGC 532
QY 288 GCTCGTTGGGCGGCTGCTCTGTTCGGTATGTAGTGGGGATCTCTGCGGATCTGT 347
Db 533 GGGCGAGCGCGCGCGCGAGAGTAGCGGAGCGCGCGCGAGCCCAATGAAGT 592
QY 348 CTTCTCTGCTCCAGCTGTTCAC 371
Db 593 CAACATGCGCACCTGCTCTCAC 616

RESULT 9
BI536071 740 bp mRNA linear EST 09-AUG-2002
LOCUS BI536071 MF01SSB cDNA Oryzias latipes CDNA clone MF01SSB002D03 3',
DEFINITION mRNA sequence.

ACCESSION BI536071
VERSION BI536071.1 GI:22194883
KEYWORDS EST.

SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
REFERENCE 1 (bases 1 to 740)
AUTHORS Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.

TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. .740
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01SSB002D03"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/db_stage="segmentation stage 20 - 25"

[illegible]

]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 66 a 162 c 101 g 94 t
ORIGIN

Query Match 6.2%; Score 39.6; DB 9; Length 423;
Best Local Similarity 47.9%; Pred. No. 7.8;
Matches 114; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 195 CAACCTCTCCCGCTGCTGGTAGCGCTCACCCACGCTCGAGCTAGGAGCCAGCGT 254
Db 83 CATGTCGTGCTGCGCGGAGGACCCGCGCTCCGAGACAGGCGACTCCCTTACGA 142
QY 255 CCCCACACAGCAATACAGACGCCACGTCGATTGCTCGTTGGGGCGGCTCTCTGTTC 314
Db 143 CCCTAGCGCTCGCGCTCCGGGCGCCCTATCTCTCTTCTGCTCTGTCCTTCTC 202
QY 315 CGCTATGATGAGGGGATCTCTCGGGATCTGTTCTTCTCGTCTCCAGCTGTTTCACT 374
Db 203 CATCAGGAGCAGCGTGACTTCAGCGAGTCCGCGAGCACCTGGCTAGACAGTTAACAA 262
QY 375 CTCGCTCCCGGCATGACCGTGCAGACTGCAATTGCTCAATCTATCCCGGCCAC 432
Db 263 CACGTCTTCCAGCCTGAGCCAGCGAGTTTGGGAAGGGGCTTCTGCGCCCCCCCAC 320

RESULT 12
CA816001
LOCUS
DEFINITION
CA12E13011VF_E04 Cabernet Sauvignon Leaf - CA12E1 Vitis vinifera
CDNA clone CA12E13011VF_E04 5', mRNA sequence.
CA816001
VERSION
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera
Vitis vinifera

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Goes da Silva, F., Lim, H., Iandolino, A., Baek, J., Jones, K., Walker, M.A. and Cook, D.R.
Transcriptional responses of Vitis vinifera to infection by the bacterial pathogen Xylella fastidiosa
Unpublished
Contact: Doug Cook
CAES Genome Facility
UC Davis Department of Plant Pathology
1 Shields Ave., Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: GTATCATGTCAGCGTACC.
Location/Qualifiers
1. .621
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA12E13011VF_E04"
/sex="hermaphrodite"
/dev_stage="Mid-season leaf material"
/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Leaf - CA12E1"
/note="Organ: Leaf; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CA12E1 is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were verified to be infected with the bacterial pathogen, Xylella fastidiosa,

FEATURES
source

based on a diagnostic assay using PCR and Xylella-specific primer pairs. The plants were asymptomatic at the time of collection, but later developed symptoms. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAGCGAGTGGCATACGGCGGG-3' and
5'-ATTCTAGAGCGGCGGCGGACATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 146 a 152 c 143 g 180 t
ORIGIN

Query Match 6.1%; Score 39; DB 14; Length 621;
Best Local Similarity 50.3%; Pred. No. 13;
Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 111 CGACTGCTCCAACTCAAGCATTTGTATGAGGAGGAGCATGATCATGCACACCCCGG 170
Db 397 CGACTTGTCTCTCATGACGGTAAAGCCAGCAGGAGCTCCGAGCTCTCCGAG 456
QY 171 GTGCGTGCCTCGTTTCGGGAGAACAACTCTTCCGCTGCTGGTAGCGCTCACCCCGC 230
Db 457 GTGCTCGAGATGCTTGTGTGCTTCTCTCTCTTCTAGTCAGGCGCTGGTCTATCTG 516
QY 231 GCTCGAGCTAGGAGCCAGCGTCCCGACGACAAATACGACGCCACGTCGATTGCT 290
Db 517 ACACCTTCGGTGGTTTCGGCAGCGTAGCGCTGACACCATCGGAAGCATAGCTGCTG 576
QY 291 CGTTGGGCGG 301
Db 577 AGTAGTGGG 587

RESULT 13
BI723734
LOCUS
DEFINITION
BI723734
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
1. .624
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant

FEATURES
source

in a medium with bubbling air containing 5% carbon dioxide"

BASE COUNT	
ORIGIN	
70 a	144 c
122 g	63 t

Query Match	6.0%;	Score 38.6;	DB 9;	Length 399;
Best Local Similarity	50.8%;	Pred. No. 14;	Indels 0;	Gaps 0;

MATCHES 26, CONSERVATION 30
 30 CTTGGATTACTGTCCTGCTGACCAATCCAGTTCGGCTTATGAGTTCGCAACGTGTC 89

Db
44 CTTTCGACCCCTGGACGGCTCGTCCATCGTCGACCACTTCGCGCTGGGCACCACTCTT
105

104 CGGCGTGTGGCCGGCGACAACTACCCGGCCGAGCAGGTGGCTGCGG 163

150 CATGATCATGCACACCCCGGGTGCCTGCGTTCGGGAGAAACACATCTCTGGCT
QY
164 CATGGGCATCTACGGTCCCGGACCGGTGTTCTGTCATTGCCCTGAAGGACGCCCCCGGCTG 223
ST

QY 210 C 210

Db 224 C 224

RESULT 15					
AV637507	linear	EST 15-DEC-2000			
LOCUS	434 bp	mRNA			
AV637507	Chlamydomonas reinhardtii				

DEFINITION	AV637507 cDNA clone HC074a01_r_5', mRNA sequence.
ACCESSION	AV637507

AV03/50711	Chlamydomonas reinhardtii
VERSION	EST.
KEYWORDS	Chlamydomonas reinhardtii
SOURCE	Chlamydomonas reinhardtii

ORGANISM
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales,
Chlamydomonadaceae; Chlamydomonas.

AUTHORS
Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Oniyama, K., Nakamura, Y. and Tabata, S.
Generation of expressed sequence tags from low-CO₂ and high-CO₂

JOURNAL
OF
MEDICAL
MICROBIOLOGY
DNA RES. 7 (5), 305-307 (2000)
adapted cells of *Chlamydomonas reinhardtii*

11089912 PUBMED
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
COMMENT

Kazusa DNA Research Institute
Kazusa
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant>

```

FEATURES
source
1. .434
/organism="Chlamydomonas reinhardtii"
location/Qualifiers

```

```

/moi_cfpd
/strain="C9"
/db_xref="taxon:3055"
/seq "H30074-01"

```

```

/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="vector: pBluescriptII SK-; Site-1: EcoRI; Site 2:
what. the cDNA library was constructed from cells cultur

```

Reaction	Temperature, °C	Time, h	Yield, %	Analysis, %
1	149	1	100	C 64.5, H 4.5, N 10.0
2	149	2	100	C 64.5, H 4.5, N 10.0
3	149	3	100	C 64.5, H 4.5, N 10.0
4	149	4	100	C 64.5, H 4.5, N 10.0
5	149	5	100	C 64.5, H 4.5, N 10.0
6	149	6	100	C 64.5, H 4.5, N 10.0
7	149	7	100	C 64.5, H 4.5, N 10.0
8	149	8	100	C 64.5, H 4.5, N 10.0
9	149	9	100	C 64.5, H 4.5, N 10.0
10	149	10	100	C 64.5, H 4.5, N 10.0
11	149	11	100	C 64.5, H 4.5, N 10.0
12	149	12	100	C 64.5, H 4.5, N 10.0
13	149	13	100	C 64.5, H 4.5, N 10.0
14	149	14	100	C 64.5, H 4.5, N 10.0
15	149	15	100	C 64.5, H 4.5, N 10.0
16	149	16	100	C 64.5, H 4.5, N 10.0
17	149	17	100	C 64.5, H 4.5, N 10.0
18	149	18	100	C 64.5, H 4.5, N 10.0
19	149	19	100	C 64.5, H 4.5, N 10.0
20	149	20	100	C 64.5, H 4.5, N 10.0
21	149	21	100	C 64.5, H 4.5, N 10.0
22	149	22	100	C 64.5, H 4.5, N 10.0
23	149	23	100	C 64.5, H 4.5, N 10.0
24	149	24	100	C 64.5, H 4.5, N 10.0
25	149	25	100	C 64.5, H 4.5, N 10.0
26	149	26	100	C 64.5, H 4.5, N 10.0
27	149	27	100	C 64.5, H 4.5, N 10.0
28	149	28	100	C 64.5, H 4.5, N 10.0
29	149	29	100	C 64.5, H 4.5, N 10.0
30	149	30	100	C 64.5, H 4.5, N 10.0
31	149	31	100	C 64.5, H 4.5, N 10.0
32	149	32	100	C 64.5, H 4.5, N 10.0
33	149	33	100	C 64.5, H 4.5, N 10.0
34	149	34	100	C 64.5, H 4.5, N 10.0
35	149	35	100	C 64.5, H 4.5, N 10.0
36	149	36	100	C 64.5, H 4.5, N 10.0
37	149	37	100	C 64.5, H 4.5, N 10.0
38	149	38	100	C 64.5, H 4.5, N 10.0
39	149	39	100	C 64.5, H 4.5, N 10.0
40	149	40	100	C 64.5, H 4.5, N 10.0
41	149	41	100	C 64.5, H 4.5, N 10.0
42	149	42	100	C 64.5, H 4.5, N 10.0
43	149	43	100	C 64.5, H 4.5, N 10.0
44	149	44	100	C 64.5, H 4.5, N 10.0
45	149	45	100	C 64.5, H 4.5, N 10.0
46	149	46	100	C 64.5, H 4.5, N 10.0
47	149	47	100	C 64.5, H 4.5, N 10.0
48	149	48	100	C 64.5, H 4.5, N 10.0
49	149	49	100	C 64.5, H 4.5, N 10.0
50	149	50	100	C 64.5, H 4.5, N 10.0
51	149	51	100	C 64.5, H 4.5, N 10.0
52	149	52	100	C 64.5, H 4.5, N 10.0
53	149	53	100	C 64.5, H 4.5, N 10.0
54	149	54	100	C 64.5, H 4.5, N 10.0
55	149	55	100	C 64.5, H 4.5, N 10.0
56	149	56	100	C 64.5, H 4.5, N 10.0
57	149	57	100	C 64.5, H 4.5, N 10.0
58	149	58	100	C 64.5, H 4.5, N 10.0
59	149	59	100	C 64.5, H 4.5, N 10.0
60	149	60	100	C 64.5, H 4.5, N 10.0
61	149	61	100	C 64.5, H 4.5, N 10.0
62	149	62	100	C 64.5, H 4.5, N 10.0
63	149			

ORIGIN
Query Match 6.0%; Score 38.6; DB 9; Length 434;

Best Local Similarity 50.8%;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0

Qy
30 CTTGGCTTACGTCGCGTGCCATCGTCGACACCACCTTCGCCGTGGGCACCATCTT

Dd
21 CTTTCGACCCCCTGGACGGCTCGTCCATCGTCGACACCACCTTCGCCGTGGGCACCATCTT

Phys. 122; 127-135; TAP + H2O2 (1, 12, 24 hr.); TAP + sorbitol (1, 2, 6, 24 hr.); TAP + Cd (1, 2, 6, 24 hr.). polyA mRNA was purified from each sample, cloned and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

Researcher	190 g	91 t	1 others
a	210 c		

6.1%;	Score 39;	DB 12;	Length 624;
ity 46.6%;	Pred. No. 13;		
	Mismatches 141;	Indels 0;	Gaps 0

CTGCTCCAACTCAAGCATTTGTGTATGAGGCAGCGGACATGATCATGCACACCCC 167

CGTGGCCCTGGCGTTGGGAGAACAACTCTTCCCGTGCTGGGTAGCGCTCACCCC 227

CAAGACCTGCCACCTGGAGGTACATCGAGTACTGGACCTGNGCAACCTGTCCAA 472

CGCAGCTAGGAACGCCAGCGTCCCCACCACGACATTCGGCCGCGGC
| | | | | | | | | | | | | | | |
CAAAAAACAACATCTTCATGATCCCCAACCCGTGTCGCGGCGCGGGCGC

TTGGGGCGGCTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGT 347

TCGTCCTCCAGCTGTTAC 371

TGGGCACCCCTGCTGCTCAC 616

399 bp mRNA linear EST 15-DEC-200

1 Chlamydomonas reinhardtii 5% CO₂ Chlamydomonas reinhardtii
1 one HC087d07_r 5', mRNA sequence.

1.1 GI:10781841
omnas reinhardtii

omonas reinhardtii; Chlorophyta; Chlorophyceae; Volvocales; Viridiplantae; Chloromonadaceae; Chlamydomonas.

Es 1 to 399)
E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohya, K.,
a, Y. and Tabata, S.

3. 7 (5), 305-307 (2000)

14
12
11: Erika Asamizu
10: Erika Asamizu
9: Erika Asamizu
8: Erika Asamizu
7: Erika Asamizu
6: Erika Asamizu
5: Erika Asamizu
4: Erika Asamizu
3: Erika Asamizu
2: Erika Asamizu
1: Erika Asamizu

DNA Research Institute
532-3, Kisarazu, Chiba 292-0812, Japan
URL: <http://www.kazusa.or.jp/en/plant/>

```
Location/Qualifiers
1. .399
/organism="Chlamydomonas reinhardtii"
```

```
/mol_type="mRNA"  
/strain="C9"  
/db_xref="taxon:3055"
```

```
/clone="HC087d07_f"
```

AP01; the CDNA library was constructed.

Qy		90	CGGATGTACCATCTCAGAACGACTGCTCCAACTCAAGCATTTGTATTGAGGCAGCGGA	149
Db		81	CGGGTGTGGCCCGCGACAAGCTGACCACAATCACCGCCGCGAGCAGGTGGCTGCCGG	140
Qy		150	CATCATATGCACACCCC CGGGTGCGTCCTCGGTCGGGAGAACAACCTTCTCCCGCTG	209
Db		141	CATGGGATCTTACGGTCCCCCGACCGGTCTTGCA TTGCCCTGAAGGACGCCCGCGCTG	200
Qy		210	C 210	
Db		201	C 201	

Search completed: December 20, 2003, 06:54:34
Job time : 1658.58 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 46.3008 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-3
Perfect score: 642
Sequence: 1 ATGCCGGTGTCTTCTC.....TACTCTTGTCTCTAATAG 642

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642	100.0	642	3	US-08-612-973-3
2	642	100.0	642	3	US-08-927-597-3
3	628.2	97.9	795	3	US-08-612-973-5
4	628.2	97.9	795	3	US-08-927-597-5
5	619.8	96.5	2082	3	US-08-612-973-47
6	619.8	96.5	2082	3	US-08-927-597-47
7	619.8	96.5	2433	3	US-08-612-973-49
8	619.8	96.5	2433	3	US-08-927-597-49
9	563.8	87.8	1539	2	US-08-470-426B-17
10	563.8	87.8	1863	2	US-08-470-426B-14
11	559	87.1	742	1	US-08-081-072-18
12	559	87.1	742	1	US-08-449-093A-18
13	559	87.1	932	1	US-08-081-072-15
14	559	87.1	932	1	US-08-449-093A-15
15	555.8	86.6	2116	3	US-08-191-160-21
16	555.8	86.6	9595	3	US-09-014-416-4
17	555.8	86.6	9599	3	US-09-014-416-6
18	544.6	84.8	9472	4	US-08-150-204B-96
19	540.8	84.2	1167	1	US-08-324-977-9
20	540.8	84.2	1167	2	US-08-384-616-9
21	540.8	84.2	1167	2	US-08-304-686A-9
22	540.8	84.2	1167	3	US-09-315-850-9
23	540.8	84.2	1499	1	US-08-324-977-3
24	540.8	84.2	1499	2	US-08-384-616-3
25	540.8	84.2	1499	2	US-08-904-686A-3
26	540.8	84.2	1499	3	US-08-315-850-3
27	540.8	84.2	6039	1	US-08-324-977-11

ALIGNMENTS

RESULT 1
US-08-612-973-3
; Sequence 3, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..636

not do MS posted

28 540.8 84.2 6039 2 US-08-384-616-11 Sequence 11, Appl
29 540.8 84.2 6039 2 US-08-904-686A-11 Sequence 11, Appl
30 540.8 84.2 6039 3 US-09-315-850-11 Sequence 11, Appl
31 540.8 84.2 9030 1 US-08-324-977-13 Sequence 13, Appl
32 540.8 84.2 9030 2 US-08-384-616-13 Sequence 13, Appl
33 540.8 84.2 9030 3 US-08-904-686A-13 Sequence 13, Appl
34 540.8 84.2 9030 3 US-09-315-850-13 Sequence 13, Appl
35 540.8 84.2 9416 1 US-08-324-977-1 Sequence 1, Appl
36 540.8 84.2 9416 2 US-08-384-616-1 Sequence 1, Appl
37 540.8 84.2 9416 2 US-08-904-686A-1 Sequence 1, Appl
38 540.8 84.2 9416 3 US-09-315-850-1 Sequence 27, Appl
39 540.8 84.2 9416 4 US-08-823-895A-27 Sequence 21, Appl
40 511.8 79.7 576 1 US-08-086-428B-21 Sequence 21, Appl
41 511.8 79.7 576 2 US-08-468-570-21 Sequence 21, Appl
42 511.8 79.7 576 2 US-08-290-665A-21 Sequence 21, Appl
43 511.8 79.7 576 4 US-08-466-601A-21 Sequence 21, Appl
44 511.8 79.7 576 5 PCT-US95-10398-21 Sequence 21, Appl
45 510.2 79.5 576 1 US-08-086-428B-17 Sequence 17, Appl

Mon Dec 22 13:28:48 2003

```
US-08-612-973-3
Query Match      100.0%; Score 642; DB 3; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.2e-159;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..636
;
; US-08-927-597-3
;
; Query Match      100.0%; Score 642; DB 3; Length 642;
; Best Local Similarity 100.0%; Pred. No. 2.2e-159;
; Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCGGTGCTTTCTCTATCTCTTGGCTTTACTGTCTGTGACCAATTCCA 60
DB 1 ATGCCCGGTGCTTTCTCTATCTCTTGGCTTTACTGTCTGTGACCAATTCCA 60

QY 61 GCTTCGCTTATGAGTGGCGACGCTGTCGGGATGTACCATGTCCAGAACACTGCTCC 120
DB 61 GCTTCGCTTATGAGTGGCGACGCTGTCGGGATGTACCATGTCCAGAACACTGCTCC 120

QY 121 AACTCAAGCAATTGTGTATAGGAGCGGACATGATCATGACACCCCGGCTGCGTCC 180
DB 121 AACTCAAGCAATTGTGTATAGGAGCGGACATGATCATGACACCCCGGCTGCGTCC 180

QY 181 TGGCTTCGGGAGAACACTCTTCCGCTGCTGGGATCTCTGGGATCTGTCTCTCGTCC 240
DB 181 TGGCTTCGGGAGAACACTCTTCCGCTGCTGGGATCTCTGGGATCTGTCTCTCGTCC 240

QY 241 AGGAACGCCAGCTCCCGACCAATACAGCGCAATACAGCGCAATGCTGCTGGGGG 300
DB 241 AGGAACGCCAGCTCCCGACCAATACAGCGCAATACAGCGCAATGCTGCTGGGGG 300

QY 301 GCTGCTCTCTGTTCCGGTATGTAGTGGGGATCTCTGGGATCTGTCTCTCGTCTCC 360
DB 301 GCTGCTCTCTGTTCCGGTATGTAGTGGGGATCTCTGGGATCTGTCTCTCGTCTCC 360

QY 361 CAGCTGTTCAACCATCTCGCTCCCGCATGAGACGGTGCAGACTGCAATTCCTCAATC 420
DB 361 CAGCTGTTCAACCATCTCGCTCCCGCATGAGACGGTGCAGACTGCAATTCCTCAATC 420

QY 421 TATCCCGGCCACATAACAGGTCAACGATGATGATGATGATGATGATGATGATGATG 480
DB 421 TATCCCGGCCACATAACAGGTCAACGATGATGATGATGATGATGATGATGATGATG 480

QY 481 ACNACGGCCCTGGTGGTATCGCAGCTGCTCGGATCCCAAGCTGCTGTTGACATGGT 540
DB 481 ACNACGGCCCTGGTGGTATCGCAGCTGCTCGGATCCCAAGCTGCTGTTGACATGGT 540

;
; RESULT 2
; US-08-927-597-3
; Sequence 3, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


QY 541 GCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCCCTACTATTCCATGTTGGGAACTGG 600
Db 541 GCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCCCTACTATTCCATGTTGGGAACTGG 600
QY 601 GCTAAGGTTTGTATGATGATGCTACTCTTTGGCTCTCTAATAG 642
Db 601 GCTAAGGTTTGTATGATGATGCTACTCTTTGGCTCTCTAATAG 642

RESULT 3
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789
US-08-612-973-5

Query Match 97.9%; Score 628.2; DB 3; Length 795;
Best Local Similarity 98.8%; Pred. No. 9.7e-156;
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TGCCCGGTGCTCTTTCTCTATCTCTCTTGGCTTTACTGTCTGTGACCATTCAG 61
Db 155 TGCCCGGTGCTCTTTCTCTATCTCTCTTGGCTTTACTGTCTGTGACCGTTCCAG 214
QY 62 CTTCCCGCTTATAGGTGCGCAAGTGTCCGGGATGTACCATGTGCACGAAGCTGTCCA 121
Db 215 CTTCCCGCTTATAGGTGCGCAAGTGTCCGGGATGTACCATGTGCACGAAGCTGTCCA 274

QY 122 ACTCAAGCATTTGTATGAGGACGCGACATGATCATGACACCCCGGGTGGCTGCCCT 181
Db 275 ACTCAAGCATTTGTATGAGGACGCGACATGATCATGACACCCCGGGTGGCTGCCCT 334
QY 182 GCGTTGGGAGAACAACTCTTCCCGTCTGGGTAGCGGTCAACCCACGCTCGCAGCTA 241
Db 335 GCGTTGGGAGAACAACTCTTCCCGTCTGGGTAGCGGTCAACCCACGCTCGCAGCTA 394
QY 242 GGAACGCCAGCGTCCCCACCAACATACAGAGCCACGCTCGATTCTCTTGGGGCGG 301
Db 395 GGAACGCCAGCGTCCCCACCAACATACAGAGCCACGCTCGATTCTCTTGGGGCGG 454
QY 302 CTGCTCTCTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
Db 455 CTGCTCTCTTCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 514
QY 362 AGCTGTTCCACCATCTCGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 421
Db 515 AGCTGTTCCACCATCTCGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 574
QY 422 ATCCCGGCCACATACAGAGTCAACGATGATGATGATGATGATGATGATGATGATGAT 481
Db 575 ATCCCGGCCACATACAGAGTCAACGATGATGATGATGATGATGATGATGATGATGAT 634
QY 482 CAACGGCCCTGTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 541
Db 635 CAACGGCCCTGTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 694
QY 542 CCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCGGCTCTTCTCTCTCTAATAG 601
Db 695 CCGGGGGCCCATTTGGGAGTCTTGGGGGCTCGCGGCTCTTCTCTCTCTAATAG 754
QY 602 CTAAGGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642
Db 755 CTAAGGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795

RESULT 4
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:

not dp

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
-08-927-597-5

Query Match	97.9%;	Score 628.2;	DB 3;	Length 795;
Best Local Similarity	98.8%;	Prod. No. 9,7e-156;		
Matches 633;	Conservative	0;	Mismatches 8;	Indels 0; Gaps 0;
2	TGCCCGGTTGCTCTTTTCTCTATCTCTTCCTCTTGGCTTTACTGTCTGTCTGACCAATTCAG	61		
155	TGCCCGGTTGCTCTTTTCTCTATCTCTTCCTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAG	214		
62	CTTTCGCTTATGAGGTGGCAACGTTGCGGATGTACCATGTCAAGAAACGACTGTCTCCA	121		
215	CTTTCGCTTATGAGGTGGCAACGTTGCGGATGTACCATGTCAAGAAACGACTGTCTCCA	274		
122	ACTCAAGCATTTGTATGAGGACGGGACATGATCATGCAACCCCGGGTTCGTGCGCT	181		
275	ACTCAAGCATTTGTATGAGGACGGGACATGATCATGCAACCCCGGGTTCGTGCGCT	334		
182	GGTTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCGCCACGCTCGCAGCTA	241		
335	GGTTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCGCCACGCTCGCAGCTA	394		
242	GGAAAGCCAGCGTCCCCACCAACACAAATACGACGCCACGTCGATTTGCTGTGGGGCGG	301		
395	GGAAAGCCAGCGTCCCCACCAACACAAATACGACGCCACGTCGATTTGCTGTGGGGCGG	454		
302	CTGCTCTCTGTTCCGTATGTACGTGGGGGATCTCTGGGATCTGTCCTCGTCTCCC	361		
455	CTGCTTTCTGTTCCGCTATGTACGTGGGGGACCTCTTGGGATCTGTCCTCGTCTCCC	514		
362	AGCTGTTTCAACCATCTCGCTTCGCCGACATGAGACGGTGCAGGACTGCATTTGCTCAATCT	421		
515	AGCTGTTTCAACCATCTCGCTTCGCCGACATGAGACGGTGCAGGACTGCATTTGCTCAATCT	574		
422	ATCCCGGCCACATTAACGAGTCAAGTATGGCTTGGGATATGATGAACTGGTCCGCTA	481		
575	ATCCCGGCCACATTAACGAGTCAAGTATGGCTTGGGATATGATGAACTGGTCCGCTA	634		
482	CAAGCGCCCTGCTGTATCGAGCTGTCTCGGATCCCAACAGCTGTCTGGAACATGGTGG	541		
635	CAAGCGCCCTGCTGTATCGAGCTGTCTCGGATCCCAACAGCTGTCTGGAACATGGTGG	694		
542	CGGGGGCCCATTTGGGAGTCTCTGGGGGCTCGCTACTATTCCATCGTGGGAACTGGG	601		
695	CGGGGGCCCATTTGGGAGTCTCTGGGGGCTCGCTACTATTCCATCGTGGGAACTGGG	754		
602	CTAAGGTTTGAATGATGATGCTACTCTTTGCTCTCTAATAG	642		
755	CTAAGGTTTGAATGATGATGCTACTCTTTGCTCTCTAATAG	795		

RESULT 5

US-08-612-973-47

: Sequence 47, Application US/08612973

Patent No. 6150134

CURRENT APPLICATION NUMBER: US/08/612,973									
FILING DATE: 11-MAR-1996									
CLASSIFICATION: 435									
ATTORNEY/AGENT INFORMATION:									
NAME: BYRNE, THOMAS E.									
REGISTRATION NUMBER: 32,205									
REFERENCE/DOCKET NUMBER: 1487-10									
TELECOMMUNICATION INFORMATION:									
TELEPHONE: (703) 816-4000									
TELEFAX: (703) 816-4100									
INFORMATION FOR SEQ ID NO: 47:									
SEQUENCE CHARACTERISTICS:									
LENGTH: 2082 base pairs									
TYPE: nucleic acid									
STRANDEDNESS: single									
TOPOLOGY: linear									
MOLECULE TYPE: cDNA									
HYPOTHETICAL: NO									
ANTI-SENSE: NO									
FEATURE:									
NAME/KEY: CDS									
LOCATION: 1..2079									
FEATURE:									
NAME/KEY: mat_peptide									
LOCATION: 1..2076									
US-08-612-973-47									
Query Match 96.5%; Score 619.8; DB 3; Length 2082;									
Best Local Similarity 98.9%; Pred. No. 1.9e-153;									
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;									
QY	2	TGCCGGTTGCTCTTCTCTATCTTCCTCTTGGCTTTACTGTCCTGTCTGCACCAATTCAG	61						
Db	155	TGCCCGGTTGCTCTTCTCTATCTTCCTCTTGGCTTTGCTGCTGTCTGACCGTTCCAG	214						
QY	62	CTTCCGGCTTATGAGTGCACAGTGTCCGGATGTACATGTACAGAACGACTGCTCCA	121						
Db	215	CTTCGGCTTATGAAGTGGCAACGCTCTCCGGATGTACCATGTACAGAACGACTGCTCCA	274						
QY	122	ACTCAAGCATGTGTATGAGGACGGACATGATCATGCACACCCCGGGTGCGTGCCT	181						
Db	275	ACTCAAGCATTTGTATGAGGACGGACATGATCATGCACACCCCGGGTGCGTGCCT	334						
QY	182	GGCTTCGGGAAACAACCTCTTCCGCTGTGGGTAGCGCTACCCCGCAGCTCCGAGCTA	241						
Db	335	GGCTTCGGGAAACAACCTCTTCCGCTGTGGGTAGCGCTACCCCGCAGCTCCGAGCTA	394						
QY	242	GGAGCCAGCGTCCCAACACGACCAATACGACCCACGCTCGATTTGCTTGGGGCGG	301						
Db	395	GGAGCCAGCGTCCCAACACGACCAATACGACCCACGCTCGATTTGCTTGGGGCGG	454						
QY	302	CTGCTCTCTGTTCGCTATGTAGTGGGGGATCTCTGGGATCTGCTTCTTCCTCGTCTCC	361						

Db 566 CTTCCGCTTATGAAGTGGCAACGCTGTCGGGATGTACCATGTGCACGAACGATGCTCCA 625
QY 122 ACTCAAGCATTTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 181
Db 626 ACTCAAGCATTTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 685
QY 182 GGTTCGGGAGAACAACTCTCCCGTGTGGTAGCGCTCACCCCGACGCTCGAGCTA 241
Db 686 GGTTCGGGAGAACAACTCTCCCGTGTGGTAGCGCTCACCCCGACGCTCGAGCTA 745
QY 242 GGAACGCCAGCGTCCCAACACAGATAGACGCGACGCTCGATTTGCTGTTGGGCGG 301
Db 746 GGAACGCCAGCGTCCCAACACAGATAGACGCGACGCTCGATTTGCTGTTGGGCGG 805
QY 302 CTGCTCTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGCTTCTCTCCCTCC 361
Db 806 CTGCTCTCTGTTCCGCTATGTAGTGGGGACCTCTGCGGATCTGCTTCTCTCCCTCC 865
QY 362 AGCTGTTACCATCTCGCTCGCGGATGAGACGCTGAGAGCTGCAATTTGCTCAATCT 421
Db 866 AGCTGTTACCATCTCGCTCGCGGATGAGACGCTGAGAGCTGCAATTTGCTCAATCT 925
QY 422 ATCCCGGCCACATAACAGCTGCTCGGATGCGGATGATGATGATGATGATGATGATGAT 481
Db 926 ATCCCGGCCACATAACAGCTGCTCGGATGCGGATGATGATGATGATGATGATGATGAT 985
QY 482 CAAAGCCCTGTTGTTATGCGAGCTGCTCGGATGCGGATGCGGATGCGGATGCGGATGCG 541
Db 986 CAAAGCCCTGTTGTTATGCGAGCTGCTCGGATGCGGATGCGGATGCGGATGCGGATGCG 1045
QY 542 CGGGGGCCCATTTGGGAGTCTCGCGGCTCGCTTACTATTCATTTGTTGGGAACTGGG 601
Db 1046 CGGGGGCCCATTTGGGAGTCTCGCGGCTCGCTTACTATTCATTTGTTGGGAACTGGG 1105
QY 602 CTAAGGTTTGTATGATGCTACTCTTTGC 632
Db 1106 CTAAGGTTTGTATGATGCTACTCTTTGC 1136

RESULT 9

US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.

REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/59-47083.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-470-426B-17

Query Match 87.8%; Score 563.8; DB 2; Length 1539;
Best Local Similarity 93.3%; Pred. No. 8,7e-139;
Matches 589; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 2 TGGCCGGTGTCTTTCTCTATCTTCTCTTGGCTTTACTGTCTGTGACCATTCAG 61
Db 506 TGGCCGGTGTCTTTCTCTATCTTCTCTTGGCTTTGTCTGTGACCATTCAG 565
QY 62 CTTCCGCTTATGAGTGGCAACGCTGTCGGGATGTACCATGTACGAACGCTGTCCA 121
Db 566 CTTCCGCTTATGAGTGGCAACGCTGTCGGGATGTACCATGTACGAACGCTGTCCA 625
QY 122 ACTCAAGCATTTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 181
Db 626 ACTCAAGCATTTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 685
QY 182 GGTTCGGGAGAACAACTCTTCCCGTGTGGTAGCGCTCACCCCGACGCTCGAGCTA 241
Db 686 GGTTCGGGAGAACAACTCTTCCCGTGTGGTAGCGCTCACCCCGACGCTCGAGCTA 745
QY 242 GGAACGCCAGCGTCCCAACACAGATAGACGCGACGCTCGATTTGCTGTTGGGCGG 301
Db 746 GGAACGCCAGCGTCCCAACACAGATAGACGCGACGCTCGATTTGCTGTTGGGCGG 805
QY 302 CTGCTCTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGCTTCTCTCTCC 361
Db 806 CTGCTCTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGCTTCTCTCTCC 865
QY 362 AGCTGTTACCATCTCGCTCGCGGATGAGACGCTGAGAGCTGCAATTTGCTCAATCT 421
Db 866 AGCTGTTACCATCTCGCTCGCGGATGAGACGCTGAGAGCTGCAATTTGCTCAATCT 925
QY 422 ATCCCGGCCACATAACAGCTGCTCGGATGCGGATGATGATGATGATGATGATGATGAT 481
Db 926 ATCCCGGCCACATAACAGCTGCTCGGATGCGGATGATGATGATGATGATGATGATGAT 985
QY 482 CAAAGCCCTGTTGTTATGCGAGCTGCTCGGATGCGGATGCGGATGCGGATGCGGATGCG 541
Db 986 CAAAGCCCTGTTGTTATGCGAGCTGCTCGGATGCGGATGCGGATGCGGATGCGGATGCG 1045
QY 542 CGGGGGCCCATTTGGGAGTCTCGCGGCTCGCTTACTATTCATTTGTTGGGAACTGGG 601
Db 1046 CGGGGGCCCATTTGGGAGTCTCGCGGCTCGCTTACTATTCATTTGTTGGGAACTGGG 1105
QY 602 CTAAGGTTTGTATGATGCTACTCTTTGC 632
Db 1106 CTAAGGTTTGTATGATGCTACTCTTTGC 1136

RESULT 10

US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS

Qy	482	CAACGGCCCTGTGGTGATCGCAGCTGCTCCGGATCCCACTGTCGTGGACATGGTGG	541
Db	1310	CAACAGCCCTAGTGGTGTCCGAGTTGCTCCGATCCCAAGCTGTCGTGGACATGGTGG	1369
Qy	542	CGGGGGCCCATTTGGGAAGTCCTGGGGGCTCGCCTACTATTCCATGGTGGGAACTGGG	601
Db	1370	CGGGGGCCCACTGGGAGTCTTGGCGGGCTTGCTCTACTATTCCATGGTAGGAACTGGG	1429
Qy	602	CTAAGGTTTTGATTGTGATGCTACTCTTTGC	632
Db	1430	CTAAGGTCCTGATTGTGGCGCTACTCTTGC	1460

RESULT 11
US-08-081-072-18
; sequence 18, Application US/08081072
; Patent No. 5641654

GENERAL INFORMATION:
 APPLICANT: No. 5641654oru MAKI, Kenjiro YAMAGUCHI, Ayumi
 APPLICANT: TOYOSHIMA, and Michinori KOHARA
 TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC
 TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS
 TITLE OF INVENTION: DIAGNOSIS
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dike, Bronstein, Roberts & Cushman
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-4280
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk, 3.50inch, 1.4mb storage
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh Classic
 SOFTWARE: Microsoft Word Version 4.0A
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/081,072
 FILING DATE: June 22, 1993
 CLASSIFICATION: 435
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 742 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to genomic RNA
 US-08-081-072-18

```

Query Match      87.1%; Score 559; DB 1; Length 742;
Best Local Similarity 92.9%; Pred. No. 1.3e-137;
                0. Mismatches 45; Indels 0; Gaps 0;
                500 Conservative

```

QY	2	TGCCGGTGTCTTTCTCTATCTTCCCTTTGGCTTTACTGTCCTGTGTGACCAATCCAG	61
Db	94	TGCCGGTGTCTTTCTCTATCTTCCCTTTGGCTTTACTGTCCTGTGTGACCAATCCAG	153
QY	62	CTTCCGCTTATGAGTGTGGCAACGTTGTCGGGATGTACCATGTACAGAACGACTGCTCCA	121
Db	154	CTTCCGCTTATGAGTGTGGCAACGTTGTCGGGATGTACCATGTACAGAACGACTGCTCCA	213
QY	122	ACTCAAGCATTTGTGTATGAGGCAGCGGACATGATCATGCACACCCCGGGTTCGTGCCCT	181
Db	214	ACTCAAGTATTGTGTATGAGGCAGCGGACATGATCATGCATACCCCGGGTTCGTGCCCT	273
QY	182	GCCTTCGGGAGAACACTCTTCCGCTGTGGTAGCGCTACACCCGCCACGGTCGCAGCTA	241
Db	274	GCCTTCGGGAGAACACTCTTCCGCTGTGGTAGCGCTCACTCCACGTTAGCGGCA	333
QY	242	GGAAACCCACGCGTCCCAACAACAATACGACGCACGTGATTTGCTCGTTGGGGCGG	301
Db	334	GGAAACCCACGCGTCCCAACAACAATACGACGGCATGTGCAATTTGCTGTTGGGGCGG	393

```

1  NUMBER OF SEQUENCES: 33
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
4  ADDRESSEE: L.L.P.
5  STREET: 1650 M Street, N.W., Suite 800
6  CITY: Washington
7  STATE: DC
8  COUNTRY: USA
9  ZIP: 20036
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Floppy disk
12 COMPUTER: IBM PC compatible
13 OPERATING SYSTEM: PC-DOS/MS-DOS
14 SOFTWARE: PatentIn Release #1.0, Version #1.30
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/470,426B
17 FILING DATE: 06-JUN-1995
18 CLASSIFICATION: 536
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: JP 2-153402
21 FILING DATE: 12-JUN-1990
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Weillacher, Robert G.
24 REGISTRATION NUMBER: 20,531
25 REFERENCE/DOCKET NUMBER: 06/59-47083.1
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (202) 659-2811
28 TELEFAX: (202) 659-1462
29 INFORMATION FOR SEQ ID NO: 14:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 1863 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: single
34 TOPOLOGY: unknown
35 MOLECULE TYPE: DNA (genomic)
36
37 PS-08-470-426B-14

```

Query Match	87.8%;	Score 563.8;	DB 2;	Length 1863;
Best Local Similarity	93.3%;	Pred. No. 9.1e-139;		
Matches 589; Conservative	0;	Mismatches 42;	Indels 0;	Gaps 0;

QY	2	TGCCCGTTTGCCTTTCTCTATCTTCCCTCTTGCGCTTTACTGCTCTGTCTGACCAATCCAG	51
DB	830	TGCCCGTTTGCCTTTCTCTATCTTCCCTCTTGCGCTTTACTGCTCTGTGACCAATCCCGAG	889
QY	62	CTTCCGCTTATGAGGTGCGCAACGTGTCGGGATGTACCATGTACGAAACGACTGCTCCA	121
DB	890	CTTCCGCTTATGAAGTGGCAACGTGTCGGGATATACCATGTACGAAACGACTGCTCCA	949
QY	122	ACTCAAGCATTTGTATGAGCGACGGACATGATCATGCAACCCCGGGTGGTGCCCT	181
DB	950	ACTCAAGCATTTGTATGAGCGACGGACATGATCATGCAATCTCCCGGGTGGTGCCCT	1009
QY	182	GCCTTCGGGAGAACACTCTTCCCGCTCTGGGTAGCGTACCCACCGCTCGCAGCTA	241
DB	1010	GCCTTCGGGAGAACACAGTCCCGTTGCTTGGGTAGCGTCACTCCACGCTCGCGCCA	1069
QY	242	GGAAGCCAGCGTCCCCACACGACAATACGACGCCAGTCGATTTGCTCGTTGGGGCGG	301
DB	1070	GGAATGCCAGGGTCCCACTACGACAAATACGACGCCACGTCGACTTGTCTGGTGGGCGG	1129
QY	302	CTGCTCTCTGTTCCGCTATGACGTGGGGATCTCTCGGATCTGCTTTCCTCGTCTCCC	361
DB	1130	CTGCTTTCTGCTCGCTATGTACGTGGGGATCTCTCGGATCTGTTTTCTCGTCTCCC	1189
QY	362	AGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGGTGCAGGACTGCAATTGCTCAATCT	421
DB	1190	AGCTGTTCAACCTTCTCGCCTCGCGGCATGAGACAGTGCAGGACTGCAATGCTCAATCT	1249
QY	422	ATCCCGGCCACATAACAGGTCAACCGTATGGCTTCGGGATATGATGATGAATGCTGCGCTA	481
DB	1250	ATCCCGGCCAATTTATCAGGTCAACCGATGGCTTCGGGATATGATGATGAAGTCACTA	1309

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 3010.09 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-5

Perfect score: 795

Sequence: 1 AGTTGGTAAGTCAATCGA.....TACTTTTGTCTCCTTAATAG 795

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.:
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	795	100.0	795	6	A48667	A48667 Sequence 5
2	795	100.0	795	6	AR157325	AR157325 Sequence
3	795	100.0	795	6	AX452754	AX452754 Sequence
4	795	100.0	795	6	AX685006	AX685006 Sequence
5	778.8	98.0	2082	6	A48709	A48709 Sequence 47
6	778.8	98.0	2082	6	AR157350	AR157350 Sequence
7	778.8	98.0	2082	6	AX452796	AX452796 Sequence
8	778.8	98.0	2082	6	AX685048	AX685048 Sequence
9	778.8	98.0	2433	6	A48711	A48711 Sequence 49
10	778.8	98.0	2433	6	AR157351	AR157351 Sequence
11	778.8	98.0	2433	6	AX452798	AX452798 Sequence
12	778.8	98.0	2433	6	AX685050	AX685050 Sequence
13	714.8	89.9	9379	14	AF207766	AF207766 Hepatitis
14	711.6	89.5	9379	14	HCVPOLYP	AJ000009 Hepatitis
15	711	89.4	9418	14	HCV132996	AJ132996 Hepatitis
16	710	89.3	9386	14	AF165056	AF165056 Hepatitis
17	708.4	89.1	9386	14	AF165055	AF165055 Hepatitis
18	708.4	89.1	9410	14	HPCK1R2	D50481 Hepatitis C
19	706.8	88.9	3296	14	AB008441	AB008441 Hepatitis
20	706.8	88.9	3296	14	AB008443	AB008443 Hepatitis
21	706.8	88.9	9379	14	AF165052	AF165052 Hepatitis
22	706.8	88.9	9379	14	AF207761	AF207761 Hepatitis
23	705.2	88.7	3296	14	AB008442	AB008442 Hepatitis
24	705.2	88.7	3296	14	AB008446	AB008446 Hepatitis
25	705.2	88.7	9379	14	AF165051	AF165051 Hepatitis
26	705.2	88.7	9598	14	AB049101	AB049101 Hepatitis
27	703.6	88.5	3296	14	AB008447	AB008447 Hepatitis
28	703.6	88.5	9547	14	AB049091	AB049091 Hepatitis
29	703.6	88.5	9605	6	AX739971	AX739971 Sequence
30	703.6	88.5	9605	14	HCJ238799	AJ238799 Hepatitis
31	703.6	88.5	11076	6	AX036252	AX036252 Sequence
32	703.6	88.5	11076	6	AX036258	AX036258 Sequence
33	703.6	88.5	11076	6	AX036260	AX036260 Sequence
34	703.6	88.5	11076	6	AX036262	AX036262 Sequence
35	702	88.3	1880	14	HPC5TRJ4	D00832 Hepatitis C
36	702	88.3	2540	6	E04260	E04260 cDNA encodi
37	702	88.3	2540	6	E04805	E04805 cDNA to 5'
38	702	88.3	2540	6	E07391	E07391 cDNA encodi
39	702	88.3	3296	14	AB008445	AB008445 Hepatitis
40	702	88.3	9033	14	HCJ238800	AJ238800 Hepatitis
41	702	88.3	9379	14	AF165048	AF165048 Hepatitis
42	702	88.3	9448	14	HPCJ483	D13558 Hepatitis C
43	700.4	88.1	1587	14	HPCCP8P1	D10687 Hepatitis C
44	700.4	88.1	1615	14	HPCNS1SPJ	M74813 Hepatitis C
45	700.4	88.1	9344	14	AB049096	AB049096 Hepatitis

ALIGNMENTS

RESULT 1
A48667
LOCUS A48667 Sequence 5 from Patent WO9604385.
DEFINITION A48667
ACCESSION A48667
VERSION A48667.1 GI:2302380
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified
REFERENCE
1 (bases 1 to 795)
Maertens,G., Bosman,F., De,M.G. and Buyse,M.
AUTHORS
TITLE
PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL
Patent: WO 9604385-A 5 15-FEB-1996;

linear PAT 07-MAR-1997

Mon Dec 22 13:28:51 2003

INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
Location/Qualifiers
1. .795
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1. .792
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03129.1"
/db_xref="GI:2302381"
/translations="MLGKVIDTLTCGFADLVGIVPLVGAFLGAAALAHGVRVLEEDG
VNTATGLPCGSFISIFLLALLSLCTPASAIEVRNVSMYHTNDCSSNSIYVEAADM
IMHTPGCVPCVRNNSRCWALPTLAARNASVPTTIRRHVDLLVGNAAFCASMYV
GDLGVSFLVSQFLTSPRRHETVQDNCISYVPGHITGRMAWDMMNNWSPFTALVVS
QLLRIPQAVDMVAGAHGVLGLAYVSMVGNWAKVLIVMLLFAP"
mat_peptide
1. .789
/product="unnamed"
BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN
Query Match 100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.2e-167; Indels 0; Gaps 0;
Matches 795; Conservative 0; Mismatches 0;
QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGTCCGGGTT 120
DB 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGTCCGGGTT 120
QY 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180
QY 181 CTCTGGCTTTCGTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTTG 240
DB 181 CTCTGGCTTTCGTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTTG 240
QY 241 TCCGGGATGACCATGTGCAAGACGCTGCTCAACTCAAGCTTCTATGAGGACGCG 300
DB 241 TCCGGGATGACCATGTGCAAGACGCTGCTCAACTCAAGCTTCTATGAGGACGCG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCGCTTCCGCTTATGAAGTCGCAACGTTG 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCGCTTCCGCTTATGAAGTCGCAACGTTG 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTGCGTCCGCTTATGAAGTCGCAACGTTTCCCGG 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTGCGTCCGCTTATGAAGTCGCAACGTTTCCCGG 420
QY 421 ATACGACGCGCTGCAATGCTGCTTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
DB 421 ATACGACGCGCTGCAATGCTGCTTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
QY 481 GGGGACCTCTCGGATCTGCTTCTCTGCTCCAGCTGTTTACCATCTCGCTCCGCTCGCGG 540
DB 481 GGGGACCTCTCGGATCTGCTTCTCTGCTCCAGCTGTTTACCATCTCGCTCCGCTCGCGG 540
QY 541 CATGAGCGGTGAGGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGT 600
DB 541 CATGAGCGGTGAGGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGT 600
QY 601 ATGGCTTGGATATGATGATGAATGCTGCGCTCAACAGGCGCTTGGTGTATGCGAGCTG 660
DB 601 ATGGCTTGGATATGATGATGAATGCTGCGCTCAACAGGCGCTTGGTGTATGCGAGCTG 660
QY 661 CTCGGATCCCAAGAGTCTGTCGACATGTCGGGGGGCCCATTTGGGGAGTCTTGGCG 720

Db 661 CTCGGATCCCAAGAGTCTGTCGTCGACATGTCGGCGGGGCCCATTTGGGAGTCTCTGGCG 720
QY 721 GGTCTCGCTACTATTCATGTCGGAACCTGGGTAAGGTTTTCATTGTGATGCTACTC 780
Db 721 GGTCTCGCTACTATTCATGTCGGAACCTGGGTAAGGTTTTCATTGTGATGCTACTC 780
QY 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795
RESULT 2
ARL57325
LOCUS ARL57325 795 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 5 from patent US 6245503.
ACCESSION ARL57325
VERSION ARL57325.1 GI:16218258
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 795)
AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 5 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..795
/organism="unknown"
BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN
Query Match 100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.2e-167; Indels 0; Gaps 0;
Matches 795; Conservative 0; Mismatches 0;
QY 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGTCCGGGTT 120
DB 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGTCCGGGTT 120
QY 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTCTCTATCTTC 180
QY 181 CTCTGGCTTTCGTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTTG 240
DB 181 CTCTGGCTTTCGTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTTG 240
QY 241 TCCGGGATGACCATGTGCAAGACGCTGCTCAACTCAAGCTTCTATGAGGACGCG 300
DB 241 TCCGGGATGACCATGTGCAAGACGCTGCTCAACTCAAGCTTCTATGAGGACGCG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCGCTTCCGCTTATGAAGTCGCAACGTTG 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCGCTTCCGCTTATGAAGTCGCAACGTTG 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTGCGTCCGCTTATGAAGTCGCAACGTTTCCCGG 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTGCGTCCGCTTATGAAGTCGCAACGTTTCCCGG 420
QY 421 ATACGACGCGCTGCAATGCTGCTTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
DB 421 ATACGACGCGCTGCAATGCTGCTTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
QY 481 GGGGACCTCTCGGATCTGCTTCTCTGCTCCAGCTGTTTACCATCTCGCTCCGCTCGCGG 540
DB 481 GGGGACCTCTCGGATCTGCTTCTCTGCTCCAGCTGTTTACCATCTCGCTCCGCTCGCGG 540
QY 541 CATGAGCGGTGAGGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGT 600
DB 541 CATGAGCGGTGAGGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGT 600
QY 601 ATGGCTTGGATATGATGATGAATGCTGCGCTCAACAGGCGCTTGGTGTATGCGAGCTG 660
DB 601 ATGGCTTGGATATGATGATGAATGCTGCGCTCAACAGGCGCTTGGTGTATGCGAGCTG 660
QY 661 CTCGGATCCCAAGAGTCTGTCGACATGTCGGGGGGCCCATTTGGGGAGTCTTGGCG 720

```
Db 541 CATGAGCGGTGAGGAGCTGCAATTTGCTCAATCTATCCCGGCCACATAAAGCGGTACCGGT 600
Qy 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTACAAACCGCCCTGGTGGTATGCGAGCTG 660
Db 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTACAAACCGCCCTGGTGGTATGCGAGCTG 660
Qy 661 CTCGGATCCCAAGCTGTGCGACATGCTGGCGGGGCCCATTTGGGGAGTCTTGGCG 720
Db 661 CTCGGATCCCAAGCTGTGCGACATGCTGGCGGGGCCCATTTGGGGAGTCTTGGCG 720
Qy 721 GGCTCGCTTACTATTCATGTTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGT 780
Db 721 GGCTCGCTTACTATTCATGTTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGT 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795

RESULT 3
AX452754
LOCUS AX452754 795 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 5 from Patent EP1211315.
ACCESSION AX452754
VERSION AX452754.1 GI:21712439
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buysse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 5 05-JUN-2002
INNOGENETICS N.V. (BE)
FEATURES
source Location/Qualifiers
1..795
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..792
/notes="unnamed protein product"
/codon_start=1
/db_xref="GI:21712440"
/mat_peptide 130 a 240 c 231 g 194 t

CDS
1..792
/translation="MLGKVIDTLTTCGADLVGYIPLVGAFLGGAARALAHGVRLVDG
VNYATGNLPGCSFISIFLLALLSCLTVPASAYEVNVSVMYHVTNDCSNSSIVYBAADM
IMHTPGCPVCRNNSRCWVALPTLAARNASVPTTIRHVDLLVGAAPFCAMTV
GDLGCVFLVSQFLTSPRRHETVQDCNCSYIPGHITGHRMWDMMNNSPTTALVVS
QLLRIPQAVVDMVAGAHGVLGLAYYSVMGNWAKVLIVMLLEAP"

BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN
Query Match 100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 3.2e-167;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CCCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGGCTCCGGTT 120
Db 61 CCCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGGCTCCGGTT 120
Qy 121 CTGAGGACGGGTGAAGTATGCAACAGGAAATTTGCCCGGTTCCTCTTCTATCTTC 180
Db 121 CTGAGGACGGGTGAAGTATGCAACAGGAAATTTGCCCGGTTCCTCTTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGCTTCCAGCTTCCCGCTTATGAAGTGCACAGCTG 240
```

```
Db 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGCTTCCAGCTTCCCGCTTATGAAGTGCACAGCTG 240
Qy 241 TCCGGGATGTACCATGTACGAAACGAGTGTCTCAAACTCAAGCATTTGTGTATGAGGACG 300
Db 241 TCCGGGATGTACCATGTACGAAACGAGTGTCTCAAACTCAAGCATTTGTGTATGAGGACG 300
Qy 301 GACATGATCATGACACCCCGGGTGGTCCCTTGGCTTCCGGAGAGAACAACTCTTCCCG 360
Db 301 GACATGATCATGACACCCCGGGTGGTCCCTTGGCTTCCGGAGAGAACAACTCTTCCCG 360
Qy 361 TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Db 361 TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Qy 421 ATACGACGCCACGCTCGATTTGCTTGGTGGGCGGCTGCTTTCTGCTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGCTCGATTTGCTTGGTGGGCGGCTGCTTTCTGCTTCCGCTATGTACGTG 480
Qy 481 GGGGACCTTGGCGGATCTGCTTCCGCTTCCGAGCTTCCAGCTTGTTCACCATCTCGCTCGCG 540
Db 481 GGGGACCTTGGCGGATCTGCTTCCGCTTCCGAGCTTCCAGCTTGTTCACCATCTCGCTCGCG 540
Qy 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGTCAACCT 600
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACAGGGTCAACCT 600
Qy 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTTCCAGCTTGTTCACCATCTCGCTCGCG 660
Db 601 ATGGCTTGGGATATGATGATGAATCGTCCGCTTCCAGCTTGTTCACCATCTCGCTCGCG 660
Qy 661 CTCGGATCCCAAGCTGTGCGAGCATGCTGGTGGGAGCTTGGTGGTGGTGGTGGTGGTGGT 720
Db 661 CTCGGATCCCAAGCTGTGCGAGCATGCTGGTGGGAGCTTGGTGGTGGTGGTGGTGGTGGT 720
Qy 721 GGCTCGCTTACTATTTCCATGTTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGTGGT 780
Db 721 GGCTCGCTTACTATTTCCATGTTGGGAACTGGGCTAAGGTTTGGTGGTGGTGGTGGTGGT 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795

RESULT 4
AX685006
LOCUS AX685006 795 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 5 from Patent WO0205548.
ACCESSION AX685006
VERSION AX685006.1 GI:29371411
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1
AUTHORS Maertens, G., Bosman, F. and Buysse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 5 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES
source Location/Qualifiers
1..795
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..792
/notes="unnamed protein product"
/codon_start=1
/db_xref="GI:29371412"
/translati="MLGKVIDTLTTCGADLVGYIPLVGAFLGGAARALAHGVRLVDG
VNYATGNLPGCSFISIFLLALLSCLTVPASAYEVNVSVMYHVTNDCSNSSIVYBAADM
```

mat_peptide		130 a	240 c	231 g	194 t	mat_peptide	
BASE COUNT						/product="unnamed"	
ORIGIN							
<p>Query Match 100.0%; Score 795; DB 6; Length 795; Best Local Similarity 100.0%; Pred. No. 3.2e-167; Indels 0; Gaps 0; Matches 795; Conservative 0; Mismatches 0;</p>							
Qy	1	ATGTGGTAAAGTTCATCGATACCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT	60				
Db	1	ATGTGGTAAAGTTCATCGATACCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT	60				
Qy	61	CCGCTCGTGGCGCCCCCTTAGGGGGCTTCAGAGGGCTTCGCGATGCGGTCGCGGTT	120				
Db	61	CCGCTCGTGGCGCCCCCTTAGGGGGCTTCAGAGGGCTTCGCGATGCGGTCGCGGTT	120				
Qy	121	CTGCAGACGCGTGAACATATGCAACAGGGAAATTTGCCGGTTCCTCTTCTATCTTC	180				
Db	121	CTGCAGACGCGTGAACATATGCAACAGGGAAATTTGCCGGTTCCTCTTCTATCTTC	180				
Qy	181	CTCTGGCTTTGCTGTCTCTGTGACCGTTCACGTTCCGCTTATGAAGTGCACAGTG	240				
Db	181	CTCTGGCTTTGCTGTCTCTGTGACCGTTCACGTTCCGCTTATGAAGTGCACAGTG	240				
Qy	241	TCCGGGATACCAATGTCACGAAAGCTGCTCCAACTCAAGCAATGTGTATGAGGACG	300				
Db	241	TCCGGGATACCAATGTCACGAAAGCTGCTCCAACTCAAGCAATGTGTATGAGGACG	300				
Qy	301	GACATGATCATGCACACCCCGGGTGTGCGCTCGGTTGGGAGAACACTCTTCCCG	360				
Db	301	GACATGATCATGCACACCCCGGGTGTGCGCTCGGTTGGGAGAACACTCTTCCCG	360				
Qy	361	TGCTGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAAGCGCAGCGTCCCAACGACA	420				
Db	361	TGCTGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAAGCGCAGCGTCCCAACGACA	420				
Qy	421	ATACGACGCGCATCGATTGCTCGTTGGGGCGGTGCTTCTGTTCCGCTATGTAGTG	480				
Db	421	ATACGACGCGCATCGATTGCTCGTTGGGGCGGTGCTTCTGTTCCGCTATGTAGTG	480				
Qy	481	GGGGACCTCTGGGATCTGCTTCCGCTTCCAGCTGTTACCATCTCGCTCGCGG	540				
Db	481	GGGGACCTCTGGGATCTGCTTCCGCTTCCAGCTGTTACCATCTCGCTCGCGG	540				
Qy	541	CATGAGACGCTGCAGGACTGCAATGCTCAATATATCCGGCGCATATACGGGTACCGT	600				
Db	541	CATGAGACGCTGCAGGACTGCAATGCTCAATATATCCGGCGCATATACGGGTACCGT	600				
Qy	601	ATGGCTTGGATATGATGATGAACTGTCGCTTACAAAGCGCTGCTGATCGAGCTG	660				
Db	601	ATGGCTTGGATATGATGATGAACTGTCGCTTACAAAGCGCTGCTGATCGAGCTG	660				
Qy	661	CTCCGATCCCAAGCTGCTGGACATGGTGGGGGGCCCACTGGGGAGTCTGGCG	720				
Db	661	CTCCGATCCCAAGCTGCTGGACATGGTGGGGGGCCCACTGGGGAGTCTGGCG	720				
Qy	721	GGTCTGGCTACTATTCCATGGTGGGACCTGGGCTAAGGTTTGTATGTGATGCTACTC	780				
Db	721	GGTCTGGCTACTATTCCATGGTGGGACCTGGGCTAAGGTTTGTATGTGATGCTACTC	780				
Qy	781	TTTGCTCCCTAATAG 795					
Db	781	TTTGCTCCCTAATAG 795					

Db	364	TGGGTAGCGCTCACCCCGAGCTCGCAGCTAGGAACGCCAGCGTCCCGACGACAAATA	423	
Qy	424	CGACGCCAGCTCGATTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGAGTGGGG	483	
Db	424	CGACGCCAGCTCGATTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGAGTGGGG	483	
Qy	484	GACCTCTGGGATCTGCTCTCTCGTCCAGCTGTTACCACTATCGCCTCGCGGCAT	543	
Db	484	GACCTCTGGGATCTGCTCTCTCGTCCAGCTGTTACCACTATCGCCTCGCGGCAT	543	
Qy	544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTCAACGGTATG	603	
Db	544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTCAACGGTATG	603	
Qy	604	GCTTGGATATGATGATGAACTGGTGCCTACAAACGGCCCTGGTGGTATGCGAGTGTCTC	663	
Db	604	GCTTGGATATGATGATGAACTGGTGCCTACAAACGGCCCTGGTGGTATGCGAGTGTCTC	663	
Qy	664	CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTTGGCGGGT	723	
Db	664	CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTTGGCGGGC	723	
Qy	724	CTCGCTTACTATTCCATGTTGGGAACTGGGCTAAAGTTTGTATGATGCTACTCTTT	783	
Db	724	CTCGCTTACTATTCCATGTTGGGAACTGGGCTAAAGTTTGTATGATGCTACTCTTT	783	
Qy	784	GC 785		
Db	784	GC 785		
RESULT 6				
LOCUS	AR157350	2082 bp	DNA linear PAT 17-OCT-2001	
DEFINITION	Sequence 47 from patent US 6245503.			
ACCESSION	AR157350			
VERSION	AR157350.1	GI:16218284		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 2082)			
AUTHORS	Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.			
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use			
JOURNAL	Patent: US 6245503-A 47 12-JUN-2001;			
FEATURES	Location/Qualifiers			
source	1..2082			
BASE COUNT	366 a	634 c	500 g 482 t	
ORIGIN	/organism="unknown"			
Query Match 98.0%; Score 778.8; DB 6; Length 2082;				
Best Local Similarity 99.7%; Pred. No. 1.3e-163;				
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Qy	4	TTGGGTAAGTCTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG	63	
Db	4	TTGGGTAAGTCTATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG	63	
Qy	64	CTCGTGGCGCCCGCTTGGGGCGCTGCGAGGCGCTGCGCATGCGCTCGCGGTCTG	123	
Db	64	CTCGTGGCGCCCGCTTGGGGCGCTGCGAGGCGCTGCGCATGCGCTCGCGGTCTG	123	
Qy	124	GAGACGGCGTGAATGCAACAGGAAATTTGCCCGGTGCTCTTCTATCTTCTCTC	183	
Db	124	GAGACGGCGTGAATGCAACAGGAAATTTGCCCGGTGCTCTTCTATCTTCTCTC	183	
Qy	184	TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTTCGCTTATGAAGTGCACAGTGTCC	243	
Db	184	TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTTCGCTTATGAAGTGCACAGTGTCC	243	
Qy	244	GGGATGTACCATGTGTCAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCGCGGAC	303	

Db	244		GGGATGTACCATGTCTCAGAACGACTGCTCTCAA	CTCAAGCATTTGTGTATGAGG	CAGCGGAC	303
Qy	304	ATGATCATGACACACCCCGGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC	363			
Db	304	ATGATCATGACACACCCCGGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC	363			
Qy	364	TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAACGCCAGCGTCCCGACGACAATA	423			
Db	364	TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAACGCCAGCGTCCCGACGACAATA	423			
Qy	424	CGAGGCCAGCTCGATTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGAGTGGGG	483			
Db	424	CGAGGCCAGCTCGATTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGAGTGGGG	483			
Qy	484	GACCTCTGCGGATCTGCTCTCTCGTCTCCACGCTGTTACCACTCTCGCCTCGCCGCGCAT	543			
Db	484	GACCTCTGCGGATCTGCTCTCTCGTCTCCACGCTGTTACCACTCTCGCCTCGCCGCGCAT	543			
Qy	544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG	603			
Db	544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG	603			
Qy	604	GCTTGGGATATGATGATGAACCTGGTGCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC	663			
Db	604	GCTTGGGATATGATGATGAACCTGGTGCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC	663			
Qy	664	CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTTGGCGGGT	723			
Db	664	CGGATCCCAAGCTGTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTTGGCGGGC	723			
Qy	724	CTCGCTTACTATTCCATGTTGGGGAACCTGGGCTAAAGTTTTGTATGATGCTACTCTTT	783			
Db	724	CTCGCTTACTATTCCATGTTGGGGAACCTGGGCTAAAGTTTTGTATGATGCTACTCTTT	783			
Qy	784	GC 785				
Db	784	GC 785				
RESULT 7						
LOCUS	AX452796	2082 bp	DNA	linear	PAT 06-JUL-2002	
DEFINITION	Sequence 47 from Patent EP1211315.					
ACCESSION	AX452796					
VERSION	AX452796.1	GI:21712481				
KEYWORDS	Hepatitis C virus					
SOURCE	Hepatitis C virus					
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.					
REFERENCE	1					
AUTHORS	Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.					
TITLE	Recombinant vectors for producing hcv envelope proteins					
JOURNAL	Patent: EP 1211315-A 47 05-JUN-2002;					
FEATURES	Innogenetics N.V. (BE)					
source	Location/Qualifiers					
	1..2082					
	/organism="Hepatitis C virus"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:11103"					
	<1..2079					
	/notes="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAD38082.1"					
	/db_xref="GI:21712482"					
CDS	/translation="NLGKVIDTLTCGFADLVGYIPLVGLPGLGAARALAHGRVLEIDG VNYATGMLPGCSFISIFLLALLSLCTVPASAYEVNYSVMYHVNDCNSIIVYEAADM IMHTPGVPCVRENSSRCWALPTPLAARNASVPTTIRHVDLLVGAARAFCSAMTV GDLGSGVFLVSQLFITSPRHEVQDCNCISYFGHITGHRMANDMMNMFYALVVS QLURIQAVVDMVAGAHGWLAGIAYSYMGNWAKVLVLMFLDGHVDFYSGVAAA SDTRGLVSLFSPGSAQIKIOLVNTNGSMHINRTALNCNDSLQDTGFFAALFKHFNSSG CPRLASCRSIDKFAQWGLPLTYTEPNSDORPCYWHYAPRPGIIVPASQVCGPVYCF					

Db 364 TGGGTAGCGCTACCCGCCAGCTCGAGCTAGGAAGCCAGCGTCCCCACGACAATA 423
Qy 424 CGAGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCTGGG 483
Db 424 CGAGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCTGGG 483
Qy 484 GACCTCTGGGATCTGCTCTCTCGTCTCCAGCTGTTACCAATCTGCGCTCGCGGCAT 543
Db 484 GACCTCTGGGATCTGCTCTCTCGTCTCCAGCTGTTACCAATCTGCGCTCGCGGCAT 543
Qy 544 GAGACGCTGAGGACTGCAATCTCAATCTATCTCCGCGCCACATAAGCGGTACCGGTATG 603
Db 544 GAGACGCTGAGGACTGCAATCTCAATCTATCTCCGCGCCACATAAGCGGTACCGGTATG 603
Qy 604 GCTTGGGATATGATGATGAATCGCTCGCTTACCAACGCGCCCTGCTGATCGAGCTGCTC 663
Db 604 GCTTGGGATATGATGATGAATCGCTCGCTTACCAACGCGCCCTGCTGATCGAGCTGCTC 663
Qy 664 CGGATCCCAAGCTGCTGTGGAATGCTGCTGCGCGGCGGCCCATTTGGGGAGTCTCTGCGGGT 723
Db 664 CGGATCCCAAGCTGCTGTGGAATGCTGCTGCGCGGCGGCCCATTTGGGGAGTCTCTGCGGGT 723
Qy 724 CTGCGCTACTATTCCATGTTGGGGAATCGGCTAAGCTTTTGTGATGCTACTCTTT 783
Db 724 CTGCGCTACTATTCCATGTTGGGGAATCGGCTAAGCTTTTGTGATGCTACTCTTT 783
Qy 784 GC 785
Db 784 GC 785

RESULT 9
A48711
LOCUS A48711 2433 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 49 from Patent WO9604385.
ACCESSION A48711
VERSION A48711.1 GI:2302424
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2433)
Maertens, G., Bosman, F., De, M.G. and Buysse, M.
PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
Patent: WO 9604385-A 49 15-FEB-1996;
INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
source
1..2433
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..2430.
/notes="unnamed protein product"
/protein_id="CAA03147.1"
/db_xref="GI:2302425"
/translation="METNPKPQRKTKRNTNRRPDQVKFPFGGGQIVGGVYLLPRRGRPL
GVARTKTSRSPORRRQPIPKARRPGRMAQPGYPWPLYNMGWAGWLLSPRG
SRPSMGPTPRRRNRLGVIDLTGCFADLVGYIPLVGAPLGGACALAHGVRLVED
GVNATGNLPGCSFSIPLALLSLCTVPASAEVRNVSGMYHVNDNCSNIIVEAAD
MIMHTPCVPCVRENNSRCWALPTLAARNASVPTTIRRHVDLLVGAAPCSAMY
VGLDGSVLVSQFLTISPRHETVQCNCSTYPGHITGRMAWDMNNWSPFTALVV
SQLRIPQAVDMVGAHWGLAGLAYYSNMGNAKVLVMLLFGVDGHTRVSGGA
ASDTRGLVSLFSPGSAQKIQLVNTNGSWHINRTALNCNDLSLOTGFFAALFYKHFNSS
GCPERLASCRSIDKFAQGWGLPTYPNSDORPCWYHAPRCGIVPASOVCGPVYC
FTPSVVGTTDRFVPTVNWGANDSDVLIINRTPRPNWGFCTWNGGTGKTCCGG
PPCNIGAGNNLTCTCDPCPKPEATYARCSPWLTFRWVHYPRMLWHYFCTVNF
TIPKRWYVGVGHRFEACNWTGRCDLEDRSELSPLLSLTTEWQLLPCSFITL
PALSTGLIHLHQNIVDQVLYGVGSVWSLVIKWEYVLLFLLLADARICACILMMLL

mat_peptide 1..2427
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
1
IAQEAALLENLVLNAAVAGAHGTLFLVFFPCAAMYIKGRLVPGAAYAPYGVWPLLL
LULADPPRAYA"
/product="unnamed"
Query Match 98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4 TTGGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 63
Db .355 TTGGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 414
Qy 64 CTGCTCGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTCTG 123
Db 415 CTGCTCGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGTCTG 474
Qy 124 GAGGACGGGCTGAACCTATGCAACAGGGAAATTTCCCGGTTGCTCTTCTATCTTCCTC 183
Db 475 GAGGACGGGCTGAACCTATGCAACAGGGAAATTTCCCGGTTGCTCTTCTATCTTCCTC 534
Qy 184 TTGGCTTTGCTGTCCTGCTGACCGCTTCAGCTTCGCTTATGAAGTGCACAGTGTCC 243
Db 535 TTGGCTTTGCTGTCCTGCTGACCGCTTCAGCTTCGCTTATGAAGTGCACAGTGTCC 594
Qy 244 GGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCGAC 303
Db 595 GGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCGAC 654
Qy 304 ATGATCATGCACACCCCGGGTGCCTGCTGCGTTGGGAGAACAACTCTTCCCGCTGC 363
Db 655 ATGATCATGCACACCCCGGGTGCCTGCTGCGTTGGGAGAACAACTCTTCCCGCTGC 714
Qy 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAAGCCAGCGCTCCCAACACGACAAATA 423
Db 715 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAAGCCAGCGCTCCCAACACGACAAATA 774
Qy 424 CGACGCGACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483
Db 775 CGACGCGACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 834
Qy 484 GACCTCTGGGATCTGCTTCTCTCGTCTCCAGCTGTTCACCATCTCGCTCGCGGCAT 543
Db 835 GACCTCTGGGATCTGCTTCTCTCGTCTCCAGCTGTTCACCATCTCGCTCGCGGCAT 894
Qy 544 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATG 603
Db 895 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATG 954
Qy 604 GCTTGGGATATGATGATGAATCGCTCGCTACAAACGGCCCTGTTGGGTATCGAGCTGCTC 663
Db 955 GCTTGGGATATGATGATGAATCGCTCGCTACAAACGGCCCTGTTGGGTATCGAGCTGCTC 1014
Qy 664 CGGATCCCAAGCTGTCGTGACATGTTGGGGGGGCCCATTTGGGAGTCTCTGCGGGGT 723
Db 1015 CGGATCCCAAGCTGTCGTGACATGTTGGGGGGGCCCATTTGGGAGTCTCTGCGGGGT 1074
Qy 724 CTGCGCTACTATTCCATGTTGGGGAACTGCGGCTAAAGCTTTTGTGATGCTACTCTTT 783
Db 1075 CTGCGCTACTATTCCATGTTGGGGAACTGCGGCTAAAGCTTTTGTGATGCTACTCTTT 1134
Qy 784 GC 785
Db 1135 GC 1136

RESULT 10
AR157351
LOCUS AR157351 2433 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 49 from patent US 6245503.
ACCESSION AR157351

VERSION AR157351.1 GI:16218285
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2433)
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buysae, M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 49 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..2433
/organism="unknown"
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
Query Match 98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTGGGTAAGTTCATCGATACCTTACATCGGCTTCGCGACCTCGTGGGTACATTCGG 63
Db 355 TTGGGTAAGTTCATCGATACCTTACATCGGCTTCGCGACCTCGTGGGTACATTCGG 414
QY 64 CTCGTCGCGCGCCCTAGAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 123
Db 415 CTCGTCGCGCGCCCTAGAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 474
QY 124 GAGGACGGCGTGAACATGATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 183
Db 475 GAGGACGGCGTGAACATGATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
Db 475 GAGGACGGCGTGAACATGATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
Db 535 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
QY 244 GGGATGATACATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTTATGAGTCCGAC 303
Db 595 GGGATGATACATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTTATGAGTCCGAC 654
QY 304 ATGATCATGACACACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 363
Db 655 ATGATCATGACACACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 714
QY 364 TGGGTAGCGCTCACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 423
Db 715 TGGGTAGCGCTCACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 774
QY 424 CGACGGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 775 CGACGGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 484 GACCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db 835 GACCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
QY 544 GAGACGGTGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
Db 895 GAGACGGTGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
QY 664 CGGATCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
Db 1015 CGGATCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
QY 724 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTGATGCTACTCTTT 783
Db 1075 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTGATGCTACTCTTT 1134
QY 784 GC 785

Db 1135 GC 1136
RESULT 11
AX452798 2433 bp DNA linear PAT 06-JUL-2002
LOCUS Sequence 49 from Patent EP1211315.
DEFINITION AX452798
ACCESSION AX452798
KEYWORDS GI:21712483
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buysae, M.-A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: Ep 1211315-A 49 05-JUN-2002;
Inogenetics N.V. (BE)
FEATURES Location/Qualifiers
source 1..2433
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
CDS 1..2430
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38083.1"
/db_xref="GI:21712484"
/translation="MSTNPFORTKENTNRPPDVKFPGGQIVGGVYLLPRRPRRL
GVRATRTKTSRSQPRRRQPIPKARREGRANWAPGYPMPLVGNMGWAGWLLSPRG
SRPSMGPTDPRRRNLGKVIDLTLCGFDALVGIPLVAGPLGGAALAHGVRLRD
GVNATGNLPGCSFISILLALLSLTVPASAEVNRVSGMTHTVNDSCNSISIVYEAAD
MIMHPGCVPCVRENNSRCWALTPDAAFNASVPTTTHRHVDLLVGLAAAPFALVY
VMDLGSFVLSQLTTPSRHETVQDCNCSIYPGTHHGMVMMNMMNSPCTSCVY
SOLLIPQAVDMVAGAHGVLAVYVSGVWNAKLVNMLLPAGDGHTRVSGAA
ASDTRGLSLFSPGSAKIQLVNINGSWHINRTALNCNDSLOTGFFAALFYKHFNFS
GPERLASCRISDFQAQGGPLTTEPNSSDQRPYCWYAPCGIWPASVCGPVCY
FPPSVVVGTTDRFGVPTYNMGANDSDVLLINNRPRGNWFGCTWNGTGTCTCGG
PCNIGAGNNTLTCTDPRFPEATYARCGSPWLTFRWVHYFPLMWHYPCVNF
TIFKVMVGVGVEHFEACNWRCDLDRSLSPLLSLTETTEWQILPCSFYTL
PALSTGLIHLHONIVDVQVLYGVGSVSLVWKVEYVLLFILLADARICAILWMLL
IAQAAALENLVNLNAAVAGAGHTLSFLVFCFAAWYIKGRILVPGAAYAFYGVNPLLL
LLLLAPPRAIA"
mat_peptide 1..2427
/product="unnamed"
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
Query Match 98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTGGGTAAGTTCATCGATACCTTACATCGGCTTCGCGACCTCGTGGGTACATTCGG 63
Db 355 TTGGGTAAGTTCATCGATACCTTACATCGGCTTCGCGACCTCGTGGGTACATTCGG 414
QY 64 CTCGTCGCGCGCCCTAGAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 123
Db 415 CTCGTCGCGCGCCCTAGAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 474
QY 124 GAGGACGGCGTGAACATGATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 183
Db 475 GAGGACGGCGTGAACATGATGCAACAGGGAATTCGCCGGTTCCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
Db 535 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
QY 244 GGGATGATACATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTTATGAGTCCGAC 303
Db 595 GGGATGATACATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTTATGAGTCCGAC 654

QY 304 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTCCCGCTGC 363
 Db 655 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTCCCGCTGC 714
 QY 364 TGGGTAGCGCTCACCCCGAGCTCGACGTAGGAACGCCAGCGTCCACACGACAA 423
 Db 715 TGGGTAGCGCTCACCCCGAGCTCGACGTAGGAACGCCAGCGTCCACACGACAA 774
 QY 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 483
 Db 775 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 834
 QY 484 GACCTCTGGCGATGCTCTCTGCTCCAGCTGTTCCACATCTCGCTCGCGCGCAT 543
 Db 835 GACCTCTGGCGATGCTCTCTGCTCCAGCTGTTCCACATCTCGCTCGCGCGCAT 894
 QY 544 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGCGGCACATAGCGGTCA 603
 Db 895 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGCGGCACATAGCGGTCA 954
 QY 604 GCTTGGGATATGATGAATGCTGCTGCTACACCGGCTGCTGCTGCTGCTGCTGCT 663
 Db 955 GCTTGGGATATGATGAATGCTGCTGCTACACCGGCTGCTGCTGCTGCTGCTGCT 1014
 QY 664 CGATATCCCAAGCTGCTGAGACATGCTGCGGGGCGCCATGCGGAGTCTGCGGGT 723
 Db 1015 CGATATCCCAAGCTGCTGAGACATGCTGCGGGGCGCCATGCGGAGTCTGCGGGT 1074
 QY 724 CTGCGCTACTATTCATGCTGGGAGTCTGCGGCTGAGGCTGAGGCTGAGGCTACT 783
 Db 1075 CTGCGCTACTATTCATGCTGGGAGTCTGCGGCTGAGGCTGAGGCTGAGGCTACT 1134
 QY 784 GC 785
 Db 1135 GC 1136

RESULT 12
 LOCUS AX685050 2433 bp DNA linear PAT 29-MAR-2003
 DEFINITION Sequence 49 from Patent WO0205548.
 ACCESSION AX685050
 VERSION AX685050.1 GI:29371455
 KEYWORDS
 SOURCE Hepatitis C virus
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1
 AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.
 TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
 JOURNAL Patent: WO 0205548-A 49 18-JUL-2002;
 INNOGENETICS N.V. (BE)
 FEATURES
 source Location/Qualifiers
 1..2433
 /organism="Hepatitis C virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:11103"
 1..2430
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD86536.1"
 /db_xref="GI:29371456"
 /translations="MSTNPKPQKTKNRRPDVKFPQGGQIVGGVLLPRRGPRL
 GVRATKTSERQPRGRQIPKARPEGRWAQPGYPWPLYNGEWMGAWLLSPRG
 SRPSWGTDPDRSRNLGRKVIDTLTCGFADLVGYIPLVGAFLGGAARALAHGVRLFD
 GNVATNLPGCCSFIFLLALLSCLTPASAYEVNRVNVGWHVNDNCSNIIYERAD
 MIMHTGCVPCVRENSRCWALTPFLARNASVPTTIRRHVDLLVGAARFCSAMY
 VGLDGSFVLVSQFLTSPRHTVQDCNCSIIYPGHITGRHMAWMMNWSPTALVV
 SQLRIPQAVWVAGAHGVLGLAYISVMGNWAKVLVVMLLFAGVDGTRVSGGAA
 ASDTRGLVSLFSPGSAKIQLVNTNGSWHINRTALNCNDSLQTFGAALFYKHFNS

GCBRLASCRSIDKFAQMGPLTYTBPNSDQRPYCMHYAPRCGIVPASPQVCGPVYC
 PTPSVVVGTTDRFGVPTYNWANDSDVLILNNRPPRNGFCWNMGTFGTCTCGG
 PCNIGAGNNLTCTCDPFRKHEATYACGSGPMLTPRCMVHYVRLWHYPCTVNF
 TIFKVMYVGGVEHFEACNWRGRCDLDRSLSPLLLSTTQWQILPQSFTTL
 PALSTGLIHLHONIVDQVLYGVGSVAVSLVIKWEYVLLFLLLADARIACULMMMLL
 IAQAEALLENLVLNNAANAGATLSFLVFFCAAWYIKGRLVPGAAYAFYGVWPLLL
 LLLALPPRAYA"
 1..2427
 mat_peptide /product="unnamed"
 BASE COUNT 434 a 745 c 714 g 540 t
 ORIGIN
 Query Match 98.0%; Score 778.8; DB 6; Length 2433;
 Best Local Similarity 99.7%; Pred. No. 1.3e-163;
 Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TTGGTAAAGTCAATCGATACCTTACATGCGGGCTTCCGCCACCTCGTGGGGTACATTC 63
 Db 355 TTGGTAAAGTCAATCGATACCTTACATGCGGGCTTCCGCCACCTCGTGGGGTACATTC 414
 QY 64 CTGCTCGGGCCCGCTTAGGGGGCGCTCCAGAGGCCCTTGGCGCATGGCGTTCGGGTCTG 123
 Db 415 CTGCTCGGGCCCGCTTAGGGGGCGCTCCAGAGGCCCTTGGCGCATGGCGTTCGGGTCTG 474
 QY 124 GAGACGCGCTGAACATATGCAACAGGGAATTTGCGGGTTCCTCTTCTATCTTCTC 183
 Db 475 GAGACGCGCTGAACATATGCAACAGGGAATTTGCGGGTTCCTCTTCTATCTTCTC 534
 QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAGTGCAGAGTCTCC 243
 Db 535 TTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAGTGCAGAGTCTCC 594
 QY 244 GGGATGTACCATGTCTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGAC 303
 Db 595 GGGATGTACCATGTCTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGAC 654
 QY 304 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTCCCGCTGC 363
 Db 655 ATGATCATGACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTCTCCCGCTGC 714
 QY 364 TGGGTAGCGCTCACCCCGAGCTCGACGTAGGAACGCCAGCGTCCACACGACAA 423
 Db 715 TGGGTAGCGCTCACCCCGAGCTCGACGTAGGAACGCCAGCGTCCACACGACAA 774
 QY 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 483
 Db 775 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGACGTGGG 834
 QY 484 GACCTCTGGCGATGCTCTCTGCTCCAGCTGTTCCACATCTCGCTCGCGCGCAT 543
 Db 835 GACCTCTGGCGATGCTCTCTGCTCCAGCTGTTCCACATCTCGCTCGCGCGCAT 894
 QY 544 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGCGGCACATAGCGGTCA 603
 Db 895 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGCGGCACATAGCGGTCA 954
 QY 604 GCTTGGGATATGATGAATGCTGCTGCTACACCGGCTGCTGCTGCTGCTGCTGCT 663
 Db 955 GCTTGGGATATGATGAATGCTGCTGCTACACCGGCTGCTGCTGCTGCTGCTGCT 1014
 QY 664 CGATATCCCAAGCTGCTGAGACATGCTGCGGGGCGCCATTGGGAGTCTGCGGGT 723
 Db 1015 CGATATCCCAAGCTGCTGAGACATGCTGCGGGGCGCCATTGGGAGTCTGCGGGT 1074
 QY 724 CTGCGCTACTATTCATGCTGGGAGTCTGCGGCTGAGGCTGAGGCTGAGGCTACT 783
 Db 1075 CTGCGCTACTATTCATGCTGGGAGTCTGCGGCTGAGGCTGAGGCTGAGGCTACT 1134
 QY 784 GC 785
 Db 1135 GC 1136

RESULT 13
AF207766
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
1 (bases 1 to 9379)
Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Marumo,F. and
Sato,C.
Characteristics of hepatitis C viral genome associated with disease
progression
Unpublished
2 (bases 1 to 9379)
Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Izumi,N. and
Sato,C.
Direct Submission
Submitted (23-NOV-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
Location/Qualifiers
1..9379
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD25"
/db_xref="taxon:11103"
/note="Isolated from a patient with liver cirrhosis and
hepatocellular carcinoma (HCC-5)"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAF65956.1"
/db_xref="GI:7650250"
/translation="MSTNPKPQKTKRNTNRNRRPDVRFPGGGQIVGGVYLLPRGRL
GYRATRSRSPGRQPIPKARRPEGRWQPGVPLYGNEGHWAGLLSPRG
SRPSGPTDPRSRNLGKVIDLTCGFADLMGYIPLVAGPLGGAALAHGVRVLD
GVNATLPGCSFIFILLASCLTI PASAEVRNAGSVYHVDLNDSSNIVYEAD
MIMTPGCVPCVENNRSRCWALTPTLAARNA SVTKTIRRHVDLLVGAAGFCAM
VGDLCGSVLSQFTFSRRHETVQDCNCSIYPGHLGSHBMDMMNMTAAVLV
SOLLTPQAVMDVAGAHGVLGAYYSYHVMGNWAKVLI VMLLPAGVDGSTRVSGTA
GRTVGLARFPTSGSQNTQLVNTGSHYHNRNALNCNDSLTQGF IALFTYHKPNS
GCPERMASCRSIDKPGQWGPITYAGDHSLDQPCYHAFPGCIPVAAQVCGPVIC
FPSPVGVTSRSGVPTNWGNSVGLLNTNRPQGNMFGCTWMNSTGTITCGG
PPCNIGGVGNNTLICPTDCFRKHEATYTKCGSPWLTPRCLVDYPPYRLMHPCTVNF
TILKVMYGVGVHRLNAACNTRGRCNLEDRSELSPILLSTTEWQILPCSFPTIL
PALSTGLIHLHQNVVDVQIDYIGISAVSVFVUKMEYVILLFALLADARCAALWMLL
IAQAEALAEKLVNAAVAGAHGILSLFVFFCAAWYIKGLVPCAALYAVXVMPILL
LILALPPLAYMDREMAASCGGAVFVGLVLLTSLSHYKVFLLARLLIWNLYLITRAEL
LOVVPPLNIRGRDAI ILLTCAHPELIFDI SKILLAILGLPTVLQAGITRVPFVR
AQLGLVCMVLRVAGHYVQAMKLAALITGVYVDHLTPLRDWAHAGLRDLAVAE
PVIFSMETALVITSLTRDKNVGEVQVSTATQSFATCVNGVCTWTVHAGSK
AYCQTKGLGCLITSLTRDKNVGEVQVSTATQSFATCVNGVCTWTVHAGSK
TLAGPKGPIITOMTVNDVDDLLVGLPAGHPTSGSKTKVPAAYAAQGYKVLVNSVATL
TRSPVFTDNSPAPVQTFQVHLHPTSGSKTKVPAAYAAQGYKVLVNSVATL
GFGAYMSKAGVDPNIRGTITTTGAPITTYTGFLADGGCGGAYDIIMCDECHS
TDSPIIGITGVLDQATAGARLVLTATPPGTVTPHPNTEALSNTGIBIPFGK
AIPLEAKGRHLLIFCHSKKDELAAGLGLNNAVAYVRGLDVSIVPVGDDVVVVA
TDALMTGFTGDLSDVDCNTCVTVDFSLDPTFTETTVPODAVSRQRRTGRG
RRGYRFTVPEERSPDSVLCCEYDAGCANVELTPAETS VRLRAYLNTPLGPVQC
DHLBEFVESVFTGLTHDAHFLSQTKQAGDNILVAYQATVCARAAQPPSMDWMKC
LIRLKPTLHGPTPLLYLRGAVQNEVILTHPIKYIMACSADLEVTISWIVGGLVA
ALAAVCLTGVSWIVGRILISGREPAVVDRELYREFDEMEECASHLPYIQOQMLAE
QFKOKALLGTASKQAEAAAPVVEKSWALEAFWAKMWNFI SGTQYLAGLSTLPGN
PALASIMPTASITSLPTQSTLLFNILGGWAAQAPPSAASAPVAGTAGAAVGS
GLKRVLDILAGYGAGVALVA FKVMGSEMPSTEDLVNLLIPALSPGALVVGWAA
ILRRHVHGEAGVQNMNRLIAPASRGNHVSPTHYVPESDAARVQTQLSLITQLLK
RLHWINDSDSTPCGSEWLWDWBI CTVLDTKTLQSKLLPRLPGVPLISQCRGYK
GYWRDGGIMHTTCCGAGIUTHVKNXGSMRIVGPRTCNTWHGTFPVNAYITGCTPSP

BASE COUNT 1853 a 2852 c 2719 g 1955 t
ORIGIN
Query Match 89.9%; Score 714.8; DB 14; Length 9379;
Best Local Similarity 94.6%; Pred. No. 2.4e-149;
Matches 740; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 4 TTGGGTAAGGTATCGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCG 63
Db 684 TTGGGTAAGGTATCGATACCTTCACATGCGGCTTCGCCGACCTCGTGGGGTACATTCG 743
QY 64 CTGCTCGCGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGCTCGGGTCTG 123
Db 744 CTGCTCGCGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGCTCGGGTCTG 803
QY 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCGCGGTTCTCTTTCTCTATCTCTC 183
Db 804 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCGCGGTTCTCTTTCTCTATCTCTC 863
QY 184 TTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGCAGCAAGTGTCT 243
Db 864 TTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGCAGCAAGTGTCT 923
QY 244 GGGATGATCATGACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCTCCCGCTGC 303
Db 924 GGGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 983
QY 304 ATGATCATGACACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCTCCCGCTGC 363
Db 984 ATGATCATGACACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCTCCCGCTGC 1043
QY 364 TGGGTAGCGCTACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCTCCCGCTGC 423
Db 1044 TGGGTAGCGCTACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCTCCCGCTGC 1103
QY 424 CGAGCGCATCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 1104 CGAGCGCATCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
QY 484 GACCTCTGCGGATCTGCT 543
Db 1164 GACCTCTGCGGATCTGCT 1223
QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTACCGGTATG 603
Db 1224 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTACCGGTATG 1283
QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 1284 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1343
QY 664 CGGATCCCAAGCGTGTGCTGAGCATGGTGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT 723
Db 1344 CGGATCCCAAGCGTGTGCTGAGCATGGTGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT 1403
QY 724 CTCGCCCTACTATTCCATCGTGGGGAACATGGGCTTAAGGTTTTGATGTGATGCTACTCTT 783

```

Db      1404  CTTGCCCTACTATTCCATGGTGGGGAACCTGGCTAAGGTCTTGATGTGATGCTACTCTTC 1463
Qy      784  GC 785
Db      .1464  GC 1465

RESULT 14
LOCUS   HCVPOLYP
DEFINITION Hepatitis C virus complete genome sequence.
ACCESSION AJ000009
VERSION   AJ000009.1
KEYWORDS complete genome; core protein; E1 protein; E2 protein; NS2 protein;
          NS3 protein; NS4a protein; NS4b protein; NS5a protein; NS5b
          protein; p7 protein; polyprotein.

SOURCE  Hepatitis C virus
ORGANISM Hepatitis C virus
          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus.

REFERENCE 1 (bases 1 to 9379)
AUTHORS Trowbridge, R. and Gowans, E.J.
TITLE Molecular cloning of an Australian isolate of hepatitis C virus
JOURNAL Arch. Virol. 143 (3), 501-511 (1998)
MEDLINE 98233263
PUBMED 9572551

REFERENCE 2 (bases 1 to 9379)
AUTHORS Trowbridge, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1997) Trowbridge R., Hepatitis Unit, Sir Albert
          Sakzewski Virus Research Centre, Royal Children's Hospital, Herston
          Road, Brisbane, Queensland Q 4029, AUSTRALIA
          Location/Qualifiers
            1..9379
              /organism="Hepatitis C virus"
              /viroid
              /mol_type="genomic RNA"
              /isolate="HCV-A"
              /db_xref="taxon:11103"
            1..9379
              /notes="genomic RNA for polyprotein precursor, complete
              genome"

misc_RNA
1..324
325..9357
/codon_start=1
/product="polyprotein"
/protein_id="CAA03854.1"
/db_xref="GI:2764398"
/db_xref="SPTREMBL:O93016"
/translations="MSTNPKPQKTKRNTNRNPDQVKFGGQIIVGGVYLLPRGPRRL
          GVRLARKTSRSPGRGRQPIPKARRPEGRMAQPGYPWPYLGNEGMWGLLSPRG
          SRPWSGPDPRRSBNLGVKIDITLTCGADLMGYIPLVGPALGGAARALAHGVRLSD
          GVNAYAGNLPGCCSFSILLGLLCLTIPASAYEVRNVSGVHVINDCSNASIYERAD
          MIVHPGCPVPCVRDNSSRCWALTPLAARNASVPTAIRRHVDLLVGAATCSAMY
          VGLDCGSFVLAQLFTFSRPHETVQDCNGSIYPGHITGHRMAMWMMNSPTAALVY
          SOLLRIPOALVDMVAGHMGVLGAYVSMVGNKAVLI VMLLFAGVDGHTHTGGVA
          QDRLRTFGFSLGPKQIKIOLVNTGSMWIRNLTALNCNDSLTGWLAAALFYTHSPNAS
          GCPERWASCHPIDEFAGGWPITYAEHSSSDQRPYCHYAPQPCGIVPASEVCGPYVC
          FTSPVVVGTTHRGVTSWGENGTDLVLLNTRPPQGNWFGCTWNGTGFTKTCGG
          FPCNIGVGNTLTCTPDCPKRHPATYTKCGSGPMLTPRCMDVDPYRLWHYPCTVNF
          TIFKRMVGVGVEHRLSAA CNWTRGERCDLEDRLSPLLLSTTEWQVLPGSFTTL
          PALSTGLIHLHQINVDQYLGIVSGVSPFAIKWEYVLLPLLLADARACALWMLL
          IAOEAALENLVNNAASVAGAGHILSFVFCFAAWYIKGRVLPGRAYALYGYWPLL
          LLALPPRAYALDREMAASCGGAVFGLVLLTSLSPHYKEFLARLILWLYQYITRAEL
          LQWIPPLNVRGDRDAILLT CVVHPPELIFDITKLILAMIGPPMWLQAVITKPYFVR
          AQGLIRACMLPVAGHYVQMAKLKAGLTSYVDHLTPLODWAHGLRLDLAVIE
          PVVFSDMETXIIITWADTAACGDIISGLPVSARRGRIILGPADEKQGWRLIAPIT
          AYSGOQRTGLLGCIIITSLTRDKNOVEQVSTATOSFLATCVNGCVTVVTHAGSK
          TLAGKPGVQMTVNTDQDLVGWPAPEASLTPTCTGSSDLVLTVEHADVIVRRRG
          DSRGSLSPRPVSYLKGSSGGLLCPSGHANGVIFRAAVCTRGVAKAVDFVFSMRTG
          MRSFVFDNSSPAVPETTFQVHLHAPTGSKSTKVPAAAYAAQYKVLVNLNPSVAATL
          GFGAYMSKAHGTDPNITGRTVTTTGTAPITYSTYKFLADGGSGGAYDIIICDECHS

```

```

TDTTTLIGITVDQAEATAGARLVVLAATPPGSVTVPHPNIEVALNSNIGTEIPFYK
AIPETIKGRHLIFCHSKCDELAAKSLGLNNAVYVRGLDVSIPRSGDVVVA
TDALMTGFTGDFSDVDCNTCVTQDFSLDPTFTIETTTPQDVAERSQRRGTG
RGIIYRFPVPGERSGMSVLCCEYDAGCAWVELYPAETSRLRAYLNTPLGLPVCQ
DHLEFESVFTGLTHIDAFLSOTKAGENFPVLYTAQATVCARAQAPPPSDQMWKC
LIRLKTPLHGPTLLYRLGAVONEVLTHTIKYIMACMSADILEVTTSTWVLVGGVLA
ALAAVCLTTSVIVGRILISGRPAIIPREVLQFDEMEECASHUPYIEQGMQLAE
QFKOKALGLLOTATQKAAAVPVVESKQALEAFWAKHMWNFISGIIYLAGSLTLPN
LAIASLMAFTASITSLPTQHTLLFNILGWAQAPLPPSAASAFVAGAGIAGAVGSI
GLGKLVLDIILAGYAGVAGVALVAFKVMGSEMPSTEDLVNLLPAILSGALVGVCA
ILRHVGPBGAGVQMMNRLIAFASRGNHVSPHYVPESDAAARVTOILSITLTOLLYK
RLHWNEDGCTSPSCGSLRDVNDWICTVLITDFKWLQSKLLPRLPGIPYSCORGKY
GVWGGIMQITTCGAQITGHVNGSMRILVGPCTCSNTWHGTFPINAYITTPCTSP
APYSRALRMAAEEVEYVTRVGDHYVTGTTDNVVKPCQCPVPAPEFTTVDGVRLLR
YAPACPLLRDEVTFOVLNVYVSQLPCEPEPDVVVVTSMITDPSHITAEAKRRL
DRGSPSLASSASQASPSLACTTRHSDPADLIEANLWROEMGNITRVESEN
KVILDSFDPLRAEEDEREVSAAEILIKSRERPRAMPVIAWPDYNPPLAESWKDPDY
VPPVHGCELPPTKAPPIPPRKRRTVLTSTVSSALAEATKTFGSSSSSAVDSGT
ATAPPOPPNDITGSESCSMPPLEGGPDPLSDSGSWSTVSEASBDDVCCSMS
YTWTGALITPCAAEESKLPINALNSLLRHHNVYATTSASQORKKVTFDLQVLD
DHYRDLVKEMKASTVKAKLSEVEACKLTPHSARSKFGYGAQVRLNLSKAVNH
RSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIVPDLGVRCEKMAVD
VSTLPOAVMGPVSGYSPQSVQSVFELVNAWKSCKPCMGAYDTRCFDSVTRESDIRV
EESIYCCDLAPEARQAISLTERLYIGPLTNSKQNCQYRRCRASGVLTSCGNTL
TCYLKASACRAAKLQDCTMLVCGDDLVICESAGTQEDAAISLVFTTEAMTRYSPPG
DPQPEYDLELITSSSNVAHDASGRVYVLTDRPTTFLARAAMFARTHTPVNSWL
GNIMYAPTILWARMILMTHFSLIAQOEKALDQIYGACYSIEPLDLPQIITQRLH
GLSASLHYSYSGEINRVASCLRKIGVPLRWRHARSVRALKLSGGRAATCGRYL
FNWAVTKLKLTPIPAASQLDLSNWFVAGVSGGDIYHLSLRAPRPFMWCLLLSVGV
GIVLLNPR"
325..897
/product="core protein"
898..1473
/product="E1 protein"
1474..2562
/product="E2 protein"
2563..2751
/product="p7 protein"
2752..3402
/product="NS2 protein"
3403..5295
/product="NS3 protein"
5296..5457
/product="NS4a protein"
5458..6240
/product="NS4b protein"
6241..7581
/product="NS5a protein"
7582..9354
/product="NS5b protein"
9358..9379
3'UTR 1993 a 2830 c 2694 g 1962 t
BASE COUNT
ORIGIN
Query Match 89.5%; Score 711.6; DB 14; Length 9379;
Best Local Similarity 94.4%; Pred. No. 1.3e-148;
Matches 738; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 4 TTGGGTAAGTCATCGATACCTTTACATGCGGCTTCCCGACCTCTGTTGGGGTACATTCG 63
Dy 679 TTGGGTAAGTCATCGATACCTTTACATGCGGCTTCCCGACCTCTGTTGGGGTACATTCG 738

Qy 64 CTGTCGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTCGGCGATGCGCTCCGGTCTG 123
Dy 739 CTGTCGGCGCCCGCTTAGGGGGCGCTGCCAGGCGCTCGGCGATGCGCTCCGGTCTG 798

Qy 124 GAGGACGCGTGAACATATGCAACAGGGAAATTTGCCCGGTTGCTCTTCTCTATCTTCTC 183
Dy 799 GAGGACGCGTGAACATATGCAACAGGGAAATTTGCCCGGTTGCTCTTCTCTATCTTCTC 858

Qy 184 TTGGCTTTGCTGCTCTGCTCTGACCGTTCAGCTTCCGCTTATGAAGTGCAGCAAGTGTCC 243
Dy 859 TTGGCTTTGCTGCTCTGCTCTGACCGTTCAGCTTCCGCTTATGAAGTGCAGCAAGTGTCC 918

```

/protein_id="CAB41950.1"		/db_xref="GI:4753719"	
/translation="MSTNPKPORKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLL GVRATKTSERSQPRGRQPIPKARPEGRWAQCYGWPPLYGNSGLGAGWLLGSPRG SRPMTPTDRRRSRNLKVIDTLTCGFAADLMGYIPLVGAFLGABALAHGVRLED GUNYATGNLPGCCFSIFLIALLSLCTIPASAYEVRNVSGVHVHTVDCNSNLSIVYETAD MIMHTGCVPCVREDNSSRCVVALTPTAARNGSVPTTAIRRHVLLVGAALFCSAMV VQDFCGSVFLVSQFLTSLPRRHETVQECNSIYPGHVTHRRMAWMMNMFPTALVV SOLLIPQAVDNVAGAHGVLAVYSMGVNWAKLLIMLLFAGVDDGGFTYRTIGG SQOAAAGLSVMSFVSQKILQINTGSGWHNRTALNCNDSLLNTGFLAALFYAHKFN SSCQPERMASCRPIDRPAQGWGITIAEPGSLDQRPYCHYAPQPCGIPVASEVCGPV YCTPTSPVVVGTDRSGVPTRWGDDNLTVDLLNNTNRPQGMNFGCTWNTGFTKTC GCPNCGIGLGNNTLICTPTDCFRKHSEATYTKGSGPMLTPTRCIVDYPRLRHHYPTCT NFTIFKVMYVGGIEHRLSAAACNWTGRNCNEDRDRSELSPULLSTTETWQILPCGFT TLTAPSTGLIHLHQNIVDVOYLVGIGSVLVSFAIKWEYIILLFLLLADARVAVLWML LLLAQAALNLVNLMAAGAHGILGFLVFFCAAMYIKGRVLPVCAAYALVSVMLP LHLLALUPPRAYANDREMAASCGAVPVGVLVLTSLPYVKLAKLIMWLOYPITRAE AHQWTPPLNVRGRDAIILLTCAVHPPELIPDITKLLAILGLPMLVAGLIRAPYF VRAQGLIRACMLVRKAGHYVQAMPKLAALTGTGYVYDHLTPLRDWAHAGRLAVF VEPVFSDMETKIITWGAADTAACGDIILGLPVSAARBGKEILLIGPADSLGQWRLLAP ITAYSQOTRGLGCIITSLTRGRKNQVGEVQVVSSTATQSFATCVNGVCWTYHGAG SKTLAGKPGKITQMYTNVDQDLVGMPPAPSGARSLTFTCTCGSSDLXILVRRHADVLPVR RGDSRGLSLPRPUSVLKSGSGGPELPCPSGHVVGIFRAAVCTRGVAKADVFPVBSME TTTSPVFTDNSSPPAPVQFVAHLHATPGSGKSTKVPAAVAAQGYKVLVLPNSVAA TLSPGAYMSKAGVDPSIRITGVRTITTCAPITYSTYKFLADGCGSGAYDIICDBE TLSFGAYMSKAGVDPSIRITGVRTITTCAPITYSTYKFLADGCGSGAYDIICDBE HSTDSTILGIVLDAETAGARLVVIAATATPPGVTVPHPNIEEVALNTEGIEIPY GKAIPETIRGRHLLIFCHSKKCDLBAQLSLGILGNANAYIRGLDVSVIPTSQDVVG VATDALMTGTFDFSDVIDCNTCTVTQVDFSLDPTFTIETTVPDQAVSRQBRGV RGRGIVHFTVTPGERPSCMFDSVLCYCDAGCAWYELTPTAETSVRURALINTPLGPV CODHLEFWESVFTGLTHIDAHFTSOTKAGDNFPYVAYQATVCARAQAQPPSWDQW KCLIRKPTLHGTPLLYRLGAVONEVTLTHPTKYIMACMSADLEVTVLTVLGVV LAALAACTTCTVSGVIVIGRITLSKPAIIPDRGVLYREDFEMBECAASHLYPIEQMOL AEQFKALGALLOTATKQAEAPVVSQKMALEAFWAKHWSPIGIGIYLAGLSTLP GNPAIASMAFTASVTSPLTQHTLILNGVAAQALAPSAASA FVAGIAGAAGV STGLKVLVDILAGVAGVAGALVAFKVMGEMPESTEDLVNLLPAILSPGALVGVVC AAILRRHVGPGEVQVWNRNIIAFASRGNHVSPTHYVPESDAAARVYQILSSLTQL LRLHQWINECDSTPCSGWLRDWDVICTVLTDFKTLQSKLLPLGFLPFIYVFFSCQPG YKGMVGRGIMOTTCPCGAQYITGVHVGSMRI VGPRTCSNTHGTFPIYATVFGVRL SPAPYFRALMRVAAEYIVTGVGDTHVVTGMTDNVKPCQVPAPEFFTEVDGVR HRYAPACKPDLREEVTFQGLNQYLVSQLPCEPEPGVTLTSMITDPSHITABTKR RLDRGPPSLASSASQSLKATCTTRHSDPADLIEANLLWRQBMGNITRVES ENKVVILDSFDFLRAEBEREISVPAEILRRSRKFPAMP IWARDPINPPLLESKDP DYVPVHVGCPLPAKAPVPPRRKRVLTSTVSSALAEALATKTGSSSESAVDS GTATAPPQPSDGGTGSDAECSMSPLEGEPPGDPDLSGDSVTSSEASESDVWCCS LDDHYWLVKEMKAKASTVKARLLSVEACKLTTPHSAKSKFGYKADKVRNLSKAIIN MSTYTGALITPCAAEESKLPINALNSLHRNNVYATTSRSASOROKVTFDRLQV HINSWVEDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLGVRVCEKAL YDVVSTLPQAVMGSSYFQYSPQVFEFLVKAWSKKNPMGFAYDTRCFDSVTENDI RVGESIYQCDLAPEARQAIRSLTERLYIIGPLTNSKQNCNGRRCSRAGSVLTTCGN TLTCYLKASAAKRAKQOCDTMLVCGDDLVVVICESAGTOEDAAASLRVFEAMTRISAP PGDPPQBYDLELITSCSNVSVAHDAKGRVYVLTDPDTPPLARAAMTARHTPVNS WLGNIIMYAPTLWNRVLMHFPFSLIAQOLEKALDCQIYGACYSIEPLDLPQIIOR HLGLSPSLHSYSPGEBINRVASCLRGVPLRVWRHARSVRKLLSRGGAATCGK YLFNWAVRTKLKLTPIPAASRLDLSGNFVAGVSGGDIYHLSLSPARPRWPMWCLLILSV GVGIIYLLPNR"			
QY	4	TTGGGTGAAGGTCATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGGTACATTCGG	63
Db	695	TTGGGTGAAGGTCATCGATACCCCTCACATGCGGCTTCGCCACCTCATGGGGTACATTCGG	754
QY	64	CTCGTCGCGGCCCCCTTAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTTCG	123
Db	755	CTCGTCGCGGCCCCCTTAGGGGGCTGCCAGGGCCCTTCGGCATGGCGTTCG	814
QY	124	GAGGACGGGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTCTC	183
Db	815	GAGGACGGGTGAATATGCAACAGGGAATCTGCCCGGTTCCTTTCTCTATCTTCTC	874
BASE COUNT 1861 a 2843 c 2716 g 1998 t			
ORIGIN			
Query Match		89.4%;	Score 711; DB 14; Length 9418;
Best Local Similarity		94.3%;	Pred. No. 1.7e-148;
Matches 738; Conservative		0; Mismatches 45; Indels	0; Gaps
QY	784	GC 785	
Db	1459	GC 1460	
RESULT 15			
HCV132996			
LOCUS			
DEFINITION			
Hepatitis C virus, complete genome, isolate HCV-AD78.			
ACCESSION			
AJ132996			
VERSION			
complete genome; polyprotein.			
KEYWORDS			
Hepatitis C virus			
SOURCE			
Hepatitis C virus			
ORGANISM			
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
REFERENCE			
1. Rispeter, K.			
AUTHORS			
JOURNAL			
Thesis (1998) Universitaetsklinikum Essen, Institut fuer Virologie			
REFERENCE			
2. (bases 1 to 9418)			
Rispeter, K.			
AUTHORS			
DIRECT SUBMISSION			
TITLE			
Submitted (07-MAR-1999) Rispeter K., Universitaetsklinikum Essen,			
JOURNAL			
Institut fuer Virologie, Hufelandstrasse 55, 45122 Essen, GERMANY			
FEATURES			
source			
1..9418			
/organism="Hepatitis C virus"			
/viral			
/mol_type="genomic RNA"			
/isolate="HCV-AD78"			
/db_xref="taxon:11103"			
/country="Germany"			
/note="genotype 1b-without entire 3'NCR"			
341..9379			
/codon_start=1			
/product="polyprotein"			
CDS			

QY 184 TTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTTGTC 243
Db |||||
875 TTGGCTTTGCTGTCCTGTCTTGACCATCCAGCTTCCGCTTATGAAGTGGCAACGTTGTC 934
QY |||||
244 GGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGCGAC 303
Db |||||
935 GGGGTGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGACGCGAC 994
QY |||||
304 ATGATCATGACACCCCGGGTGGTCCCTGCGTTCGGAGAACAACTCTTCCCGCTGC 363
Db |||||
995 ATGATCATGACACCCCGGGTGGTCCCTGCGTTCGGAGGATAAATCCTCTCGCTGC 1054
QY |||||
364 TGGGTAGCGCTCACCCCGACGCTCGACGCTAGGACGCCAGCGTCCCCACACGACAATA 423
Db |||||
1055 TGGGTAGCGCTCACCCCGACGCTCGCGGCTAGGAAATGGCAGTGTCCCCACACGCGGATA 1114
QY |||||
424 CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGGG 483
Db |||||
1115 CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGG 1174
QY |||||
484 GACCTCTGGCGATGCTGCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCCGGCAT 543
Db |||||
1175 GATTTTGGCGATGCTGTTCTCTCGTCTCCAGCTGTTCAACCTCTCTCTCGCCGGCAT 1234
QY |||||
544 GAGACGCTGACGACTGCAATTGCTCAATCTATCCCGGCCACACATAACGGGTACCGTATG 603
Db |||||
1235 GAGACAGTACAGGAATGTAATTGCTCAATTTACCCCGGCCACGTCACAGGTACCCGATG 1294
QY |||||
604 GCTTGGGATATGATGAATGCTGTCGCTACAAACGGCCCTGTTGGTATCGCAGCTGCTC 663
Db |||||
1295 GCTTGGGATATGATGAATGCTGTCCTACAAACGGCCCTAGTGGTATCGCAGCTACTC 1354
QY |||||
664 CGGATCCCAACAAGCTGCTGAGACATGTTGGCGGGGCCCAATTGGGGAGTCTCTGGCGGGT 723
Db |||||
1355 CGGATCCCAACAAGCTGCTGAGACATGTTGGCGGGGCCCACTGGGGCGTCTCTGGCGGGC 1414
QY |||||
724 CTGCGCTACTATTCCATGGTGGGGAATGGGCTAAGGTTTTGATTGTGATGCTACTCTTT 783
Db |||||
1415 CTCGCTACTATTTCGATGGTGGGGAATGGGCTAAGCTTTTGAATTGTGATGCTACTCTTT 1474
QY 784 GCT 786
Db 1475 GCT 1477

Search completed: December 20, 2003, 02:01:50

Job time : 3012.09 secs

THIS PAGE BLANK (USPTO)

XX
PT
purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

proteins - in presence of di: sulphide bond cleavage agent, to
 produce proteins suitable for direct use in vaccines or diagnostic
 assays of HCV
 Claim 23; Fig 21; 146pp; English.
 AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 and E2 protein coding sequence constructs. These sequences are included
 in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 The recombinant proteins can then be isolated using a method of the
 invention. In the method, the envelope proteins are purified by
 carrying out a disulphide bond cleavage, or a reduction step with a
 disulphide bond cleavage agent, after lysis of recombinant host cells.
 The constructs containing the purified HCV envelope proteins can be used
 for vaccinating humans against HCV, for in vitro detection of HCV
 antibodies in a sample, and in a serotyping assay for detecting one or
 more serological types of HCV present in a biological sample. The
 constructs can also be immobilised on a solid substrate and incorporated
 into a reversed phase hybridisation assay for determining the presence or
 the genotype of HCV. The new purification method preserves the
 conformation of the recombinantly expressed E1, E2 and E1/E2, and
 eliminates contaminating proteins. Antigens isolated using this method
 are more reactive with human sera than those isolated by known
 techniques.
 Query Match 100.0%; Score 795; DB 17; Length 795;
 Best Local Similarity 100.0%; Pred. No. 3.6e-204; Indels 0; Gaps 0;
 Matches 795; Conservative 0; Mismatches 0;
 1 ATGTTGGGTAAAGTTCATGATACCTTACATGCGCTTCGCCGACCTCGTGGGTACATT 60
 1 ATGTTGGGTAAAGTTCATGATACCTTACATGCGCTTCGCCGACCTCGTGGGTACATT 60
 61 CCGCTCGTGGCGCCCTAGGGGCGCTGCAGGCGCTGCGCATGGCGTCCGGTT 120
 61 CCGCTCGTGGCGCCCTAGGGGCGCTGCAGGCGCTGCGCATGGCGTCCGGTT 120
 121 CTGGAGCGCGGTGAACATATCAACAGGGAATTTGCCCGGTGCTTTCTATCTTC 180
 121 CTGGAGCGCGGTGAACATATCAACAGGGAATTTGCCCGGTGCTTTCTATCTTC 180
 181 CTCTGGCTTTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG 240
 181 CTCTGGCTTTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG 240
 241 TCCGGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGCAGCG 300
 241 TCCGGGATGATACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGCAGCG 300
 301 GACATGATCATGCACACCCCGGTGCGTCCCTGCTGGGAGAACACTTCCCGC 360
 301 GACATGATCATGCACACCCCGGTGCGTCCCTGCTGGGAGAACACTTCCCGC 360
 361 TGCTGGGTAGCGTCAACCCCGGTGCGTCCCTGCTGGGAGAACACTTCCCGC 420
 361 TGCTGGGTAGCGTCAACCCCGGTGCGTCCCTGCTGGGAGAACACTTCCCGC 420
 421 ATACGACCGCATGCTGCTGCTGCTGGGCGGCTGCTTTGTTGCTATGATGCTG 480
 421 ATACGACCGCATGCTGCTGCTGCTGGGCGGCTGCTTTGTTGCTATGATGCTG 480
 481 GGGGACCTCTGGGATCTGCTTCTGCTCCGAGCTGTTACCATCTCGCTCCCGC 540
 481 GGGGACCTCTGGGATCTGCTTCTGCTCCGAGCTGTTACCATCTCGCTCCCGC 540
 541 CATGAGCGGTGAGGACTGCAATTGCTATCTATCCCGCCACATAACGGGTACCGT 600
 541 CATGAGCGGTGAGGACTGCAATTGCTATCTATCCCGCCACATAACGGGTACCGT 600
 601 ATGCTTGGGTAAAGTTCATGATACCTTACATGCGCTTCGCCGACCTCGTGGGTACATT 660

601 ATGCTTGGGTAAAGTTCATGATACCTTACATGCGCTTCGCCGACCTCGTGGGTACATT 660
 661 CTCGGATCCCAAGCTGCTGCGGACATGCTGCGGCGGCCCATTTGGGAGTCTCTGGCG 720
 661 CTCGGATCCCAAGCTGCTGCGGACATGCTGCGGCGGCCCATTTGGGAGTCTCTGGCG 720
 721 GGTCTCGCTACTATTCATGCTGGGAACTGGGCTAAAGTTTTGATTTGATGCTACTC 780
 721 GGTCTCGCTACTATTCATGCTGGGAACTGGGCTAAAGTTTTGATTTGATGCTACTC 780
 781 TTTGCTCCCTAATAG 795
 781 TTTGCTCCCTAATAG 795
 RESULT 2
 AAL48914
 ID AAL48914 standard; DNA; 795 BP.
 XX AAL48914;
 AC AAL48914;
 XX 24-OCT-2002 (first entry)
 DT Hepatitis C virus clone HCC10A E1 protein coding sequence.
 DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 KW Hepatitis C virus.
 OS WO200255548-A2.
 PN 18-JUL-2002.
 PD 11-JAN-2002; 2002WO-BP00219.
 PF 11-JAN-2001; 2001US-260999P.
 PR 30-AUG-2001; 2001US-315768P.
 XX (INNO-) INNOGENETICS NV.
 XX Maertens G, Bosman F, Buyse M;
 PI WPI; 2002-599657/64.
 DR P-PSDB; AAO18661.
 DR New therapeutic vaccine compositions comprising at least one purified
 recombinant hepatitis C virus (HCV) single or specific oligomeric
 recombinant envelope protein E1 or E2, useful for immunizing humans
 from HCV infection -
 XX Example 2; Page 161-162; 243pp; English.
 The present invention relates to new therapeutic vaccine compositions for
 inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 composition containing at least one purified recombinant HCV single or
 specific oligomeric recombinant envelope proteins selected from an E1 and
 an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 useful for inducing HCV-specific antibodies or for immunising humans
 against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 vaccines or therapeutics, in HCV screening and confirmatory antibody
 tests, for raising antibodies, in the preparation of medicament, and for
 in vitro monitoring of HCV disease or prognosing the response to
 treatment of patients suffering from HCV infection. The present sequence
 is a coding sequence described in the exemplification of the invention.
 XX Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;
 SQ
 Query Match 100.0%; Score 795; DB 24; Length 795;
 Best Local Similarity 100.0%; Pred. No. 3.6e-204;
 Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGTTGGGTAAAGTTCATGATACCTTACATGCGCTTCGCCGACCTCGTGGGTACATT 60

```
Db 1 ATGTGGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60
Qy 61 CCCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTTGGGCATGCGTCCGGGTT 120
Db 61 CCCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTTGGGCATGCGTCCGGGTT 120
Qy 121 CTGGAGACGGGTGAATGATGCAACAGGGAATTTGCGGGTTCCTCTTCTATCTTC 180
Db 121 CTGGAGACGGGTGAATGATGCAACAGGGAATTTGCGGGTTCCTCTTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGTG 240
Db 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGTG 240
Qy 241 TCCGGGATGATACCATGTCACGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300
Db 241 TCCGGGATGATACCATGTCACGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300
Qy 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCGCTTGGGAGAACAACTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCGCTTGGGAGAACAACTTCCCGC 360
Qy 361 TGTGGGTAGCGCTCACCCACGCTCGCAGCTAGGAAAGCGCGTCCCGACCAAGCA 420
Db 361 TGTGGGTAGCGCTCACCCACGCTCGCAGCTAGGAAAGCGCGTCCCGACCAAGCA 420
Qy 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
Qy 481 GGGGACCTCTCGGATCTGCTTCTCCTGCTCTCCAGCTGTTCAACATCTCGGCTCGCGG 540
Db 481 GGGGACCTCTCGGATCTGCTTCTCCTGCTCTCCAGCTGTTCAACATCTCGGCTCGCGG 540
Qy 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGCCACATACGGTCAACGT 600
Db 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGCCACATACGGTCAACGT 600
Qy 601 ATGGCTTGGGATATGATGAACTGCTGCTTACAAAGCGGCTTGGGTATCGACGCTG 660
Db 601 ATGGCTTGGGATATGATGAACTGCTGCTTACAAAGCGGCTTGGGTATCGACGCTG 660
Qy 661 CTCGGATCCCAAGCTGCTGGAACATGCTGGCGGGGCGCCATTCGGGAGTCTTGCGG 720
Db 661 CTCGGATCCCAAGCTGCTGGAACATGCTGGCGGGGCGCCATTCGGGAGTCTTGCGG 720
Qy 721 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGATGCTACTC 780
Db 721 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGATGCTACTC 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795
```

RESULT 3

AA148939
ID AA148939 standard; DNA; 2082 BP.

AC AA148939;

XX 24-OCT-2002 (first entry)

XX Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.

XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;

XX virucide; immunostimulant; vaccine; ds.

OS Hepatitis C virus.

XX W020025548-A2.

PN

XX

18-JUL-2002.

XX 11-JAN-2002; 2002WO-EP00219.

XX 11-JAN-2001; 2001US-260699P.

XX 30-AUG-2001; 2001US-315768P.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman F, Buyse M;

XX WPI; 2002-599657/64.

XX P-PSDB; AAO18678.

New therapeutic vaccine compositions comprising at least one purified recombinant hepatitis C virus (HCV) single or specific oligomeric recombinant envelope protein E1 or E2, useful for immunizing humans from HCV infection

XX Example 2; Page 206-209; 243pp; English.

The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunizing humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence is a coding sequence described in the exemplification of the invention.

XX Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

Query Match 98.0%; Score 778.8; DB 24; Length 2082;

Best Local Similarity 99.7%; Pred. No. 1.1e-199;

Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TTGGTAAAGTCAATCGATACCTTACATGCGGCTTCCGACCTCGTGGGTACATTCG 63

Db 4 TTGGTAAAGTCAATCGATACCTTACATGCGGCTTCCGACCTCGTGGGTACATTCG 63

Qy 64 CTGCTCGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTTCG 123

Db 64 CTGCTCGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTTCG 123

Qy 124 GAGGACGGGTGAACATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTC 183

Db 124 GAGGACGGGTGAACATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTC 183

Qy 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGT 243

Db 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGT 243

Qy 244 GGGATGTACCATGTCAACAGCACTGCTCAACCTCAAGCATTTGTATGAGGACGGAC 303

Db 244 GGGATGTACCATGTCAACAGCACTGCTCAACCTCAAGCATTTGTATGAGGACGGAC 303

Qy 304 ATGATCATGACACCCCGGGTGGTGGCGTTCGCGCTTCCGGAGAACAACTCTTCCG 363

Db 304 ATGATCATGACACCCCGGGTGGTGGCGTTCGCGCTTCCGGAGAACAACTCTTCCG 363

Qy 364 TGGGTAGCGCTACCCCGGCTCGAGCTAGGACCGGCGTCCCGACACACATA 423

Db 364 TGGGTAGCGCTACCCCGGCTCGAGCTAGGACCGGCGTCCCGACACACATA 423

Qy 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTTCTTCTGTTCCGCTATGATGCTGG 483

Db 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTTCTTCTGTTCCGCTATGATGCTGG 483

Qy 484 GACCTCTGCGGATCTGCTCTCTCCGCTGCTCCAGCTGTTCAACATCTCGCCTCGCG 543

Db 484 GACCTCTGCGGATCTGCTCTCTCCGCTGCTCCAGCTGTTCAACATCTCGCCTCGCG 543

AC AAT12974;
 AC 25-SEP-1996 (first entry)
 DT HCV E1 construct HCC166.
 XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX Hepatitis C virus.
 OS WO9604385-A2.
 XX 15-FEB-1996.
 XX 31-JUL-1995; 95WO-EP03031.
 XX 29-JUL-1994; 94EP-0870132.
 XX (INNO-) INNOGENETICS NV.
 PA Bosman F, Buyse M, De Martynoff G, Maertens G;
 PI WPI; 1996-129401/13.
 XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of di: sulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX Claim 23; Fig 21; 146pp; English.
 XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the
 CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.
 XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
 SQ
 Query Match 98.0%; Score 778.8; DB 17; Length 2433;
 Best Local Similarity 99.7%; Pred. No. 1.2e-199;
 Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TTGGTAAAGTCATCGATACCTTACATGGGGCTTCGCCGACCTCGTGGGGTACATTCCG 63
 DB 355 TTGGTAAAGTCATCGATACCTTACATGGGGCTTCGCCGACCTCGTGGGGTACATTCCG 414
 QY 64 CTGTCGGGCCCCCTAGGGGGCGCTGCCAGGCGCTGGCGATGGCGTCCGGGTTCTG 123
 DB 415 CTGTCGGGCCCCCTAGGGGGCGCTGCCAGGCGCTGGCGATGGCGTCCGGGTTCTG 474
 QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTCTC 183
 DB 475 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTCTC 534
 QY 184 TTGGCTTGTCTGCTCTGACCGTTCCAGCTTATGAAGTGGCGCAAGTGTCC 243
 DB
 535 TTGGCTTGTCTGCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGTGTCC 594
 QY 244 GGGATGATACATGTCACGAACGACTGCTCAACTCAAGCAATTGTGTATGAGGAGCGGAC 303
 DB 595 GGGATGATACATGTCACGAACGACTGCTCAACTCAAGCAATTGTGTATGAGGAGCGGAC 654
 QY 304 ATGATCATGACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363
 DB 655 ATGATCATGACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 714
 QY 364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAGCCAGCGCTCCCGCCACACCAATA 423
 DB 715 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAGCCAGCGCTCCCGCCACACCAATA 774
 QY 424 CGACGCCACGTCGATTTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
 DB 775 CGACGCCACGTCGATTTGCTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 834
 QY 484 GACCTCTGGCGATCTGCTTCTCTCGTCTCCAGCTGTTTACCATCTCGCTCGCCCTCGCCGCGCAT 543
 DB 835 GACCTCTGGCGATCTGCTTCTCTCGTCTCCAGCTGTTTACCATCTCGCTCGCCCTCGCCGCGCAT 894
 QY 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 603
 DB 895 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 954
 QY 604 GCTTGGGATATGATGATGAATGCTGCTCGCTTCAACCGGCCCTGGTGGTATCGCAGCTGCTC 663
 DB 955 GCTTGGGATATGATGATGAATGCTGCTCGCTTCAACCGGCCCTGGTGGTATCGCAGCTGCTC 1014
 QY 664 CGGATCCCAACAGCTGCTGCTGACATGGTGGGGGGGCCCATTTGGGGAGTCTTGGCGGGGT 723
 DB 1015 CGGATCCCAACAGCTGCTGCTGACATGGTGGGGGGGCCCATTTGGGGAGTCTTGGCGGGGT 1074
 QY 724 CTGCGCTTACTATTCCATGGTGGGGAACTGGGCTAAGGTTTTTGTATGATGCTACTCTTT 783
 DB 1075 CTGCGCTTACTATTCCATGGTGGGGAACTGGGCTAAGGTTTTTGTATGATGCTACTCTTT 1134
 QY 784 GC 785
 DB 1135 GC 1136
 RESULT 6
 AAL48940
 ID AAL48940 standard; DNA; 2434 BP.
 XX AAL48940;
 AC AAL48940;
 XX 24-OCT-2002 (first entry)
 DT Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
 XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 XX Hepatitis C virus.
 OS WO200255548-A2.
 XX 18-JUL-2002.
 XX 11-JAN-2002; 2002WO-EP00219.
 XX 11-JAN-2001; 2001US-260699P.
 PR 30-AUG-2001; 2001US-315768P.
 XX (INNO-) INNOGENETICS NV.
 PA Maertens G, Bosman F, Buyse M;
 PI WPI; 2002-599657/64.
 DR P-PSDB; AAO18679.

XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection -
Dn

XX
Example 2: page 212-215: 243pp; English.

The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV) specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunising humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence

XX 2434 BP: 434 A: 745 C: 714 G: 541 T: 0 other;

Query Match	96.68;	Score 767.8;	DB 24;	Length 2434;
Beat Local Similarity	99.68;	Pred. No. 1.1e-196;		
Matches 780;	Conservative	0;	Mismatches 2;	Indels 1;
QY	4	TTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATCCG	63	
Db	355	TTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATCCG	414	
QY	64	CTCGTCGGCGCCCTTAGGGGGCGTGCAGAGGCCCTGGCGATGGGCTCGGGTTCG	123	
Db	415	CTCGTCGGCGCCCTTAGGGGGCGTGCAGAGGCCCTGGCGATGGGCTCGGGTTCG	474	
QY	124	GAGGACGGCGTGAACATATGCAACAGGAAATTGCCGGTTGCTTTCTCTATCTCTC	183	
Db	475	GAGGACGGCGTGAACATATGCAACAGGAAATTGCCGGTTGCTTTCTCTATCTCTC	534	
QY	184	TTGGCTTTGCTGTCGTCGTCGTCGCTTCAGACTTCGGCTTATGAAAGTCGCAAGTGTC	242	
Db	535	TTGGCTTTGCTGTCGTCGTCGCTTCAGACTTCAGCTTCGCTTATGAAAGTCGCAAGTGTC	594	
QY	243	CGGGATGTACCATGTACAGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCGA	302	
Db	595	CGGGATGTACCATGTACAGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGCGA	654	
QY	303	CATGATCATGCAACCCCGGGTCGGTCGCCCTGCGTTTCGGAGAAACAACTCTTCCCGCTG	362	
Db	655	CATGATCATGCAACCCCGGGTCGGTCGCCCTGCGTTTCGGAGAAACAACTCTTCCCGCTG	714	
QY	363	CTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAGCCAGGCTCCCAACGACGCAAT	422	
Db	715	CTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAGCCAGGCTCCCAACGACGCAAT	774	
QY	423	ACGACGCCAGTCGATTGCTCGTTGGGGCGCTTTCGTTCCGCTATGTACGTGGG	482	
Db	775	ACGACGCCAGTCGATTGCTCGTTGGGGCGCTTTCGTTCCGCTATGTACGTGGG	834	
QY	483	GGACCTCTGGGGATCTGTCTTCCTGCTCCAGCTGTTCAACATTCGCTCGCCGCA	542	
Db	835	GGACCTCTGGGGATCTGTCTTCCTGCTTCGAGCTTTCACCATCTCGCCTCGCCGCA	894	
QY	543	TGAGACGGTCAGGACTGCAATTGCTCAATCTATCCCGCCACATACCGGTCACCGTAT	602	
Db	895	TGAGACGGTCAGGACTGCAATTGCTCAATCTATCCCGCCACATACCGGTCACCGTAT	954	
QY	603	GGCTTTGGGATATGATGATGAATGAACTGGTCGCTTACAAAGGCCCTGGTGGTATCGGAGTCT	662	
Db	955	GGCTTTGGGATATGATGATGAATGAACTGGTCGCTTACAAAGGCCCTGGTGGTATCGGAGTCT	1014	
QY	663	CCGGAATCCCAACAGCTCTGCTGACATGTTGGGGGGCCCATTCGGGAGTCTCTGGGGGG	722	

CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is the HCV replicon Con 1, used as a basis for
CC the adaptive mutations of the invention.

XX
SQ Sequence: 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGTACATCCG 63
DB |||||
QY 64 CTGTCGCGCCCCCTAGGGCGCTGCCAGGCGCTTGGCGCATGGCGGTCTCTG 123
DB |||||
QY 756 CTGTCGCGCCCCCTAGGGCGCTGCCAGGCGCTTGGCGCATGGCGGTCTCTG 815
DB |||||
QY 124 GAGACGCGGTGAATCATATGACAGGGAATTTGCGCGGTGCTCTTCTATCTTCTC 183
DB |||||
QY 816 GAGACGCGGTGAATCATATGACAGGGAATTTGCGCGGTGCTCTTCTATCTTCTC 875
DB |||||
QY 184 TTGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGACGTCTCC 243
DB |||||
QY 876 TTGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGACGTATCC 935
DB |||||
QY 244 GGGATGTACCATGTTCACGACGACTGCTCACAAGCATTTGTATGAGGACGCGGAC 303
DB |||||
QY 936 GGAGTGTACCATGTTCACGACGACTGCTCACAAGCATTTGTATGAGGACGCGGAC 995
DB |||||
QY 304 ATGATCATGACACCCCGGGTGGTGGTCCCTGCTCGGAGAACAACTCTTCCCGTGC 363
DB |||||
QY 996 ATGATCATGACACCCCGGGTGGTGGTCCCTGCTCGGAGAACAACTCTTCCCGTGC 1055
DB |||||
QY 364 TGGTAGCGCTACCCCGGCTCGAGCTAGGACGCCAGCGTCCCGACACACATA 423
DB |||||
QY 1056 TGGTAGCGCTACCTCCACGCTCGCGGCCAGGAACGCTAGCGTCCCGACACATA 1115
DB |||||
QY 424 CGAGCCACGCTCGATTTGCTGTTGGGCGGCTCTTCTGTTCCGCTATGTAGTGGG 483
DB |||||
QY 1116 CGAGCCATGTGCTGCTGTTGGGCGGCTGCTCTGCTCGCTATGTAGTGGGA 1175
DB |||||
QY 484 GACTCTGCGGATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
DB |||||
QY 1176 GATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1235
DB |||||
QY 544 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGTCACTG 603
DB |||||
QY 1236 GAGACGCTGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGTCACTG 1295
DB |||||
QY 604 GCTTGGGATGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
DB |||||
QY 1296 GCTTGGGATGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1355
DB |||||
QY 664 CGATCCACAGCTGCTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
DB |||||
QY 1356 CGATCCACAGCTGCTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1415
DB |||||
QY 724 CTGCTCTACTATTCATGCTGGGAATGCTGGCTGCTGGCTGCTGGCTGCTGGCTG 783
DB |||||
QY 1416 CTGCTCTACTATTCATGCTGGGAATGCTGGCTGCTGGCTGCTGGCTGCTGGCTG 1475
DB |||||
QY 784 GC 785
DB ||
QY 1476 GC 1477

RESULT 8
ABK91424
ID ABK91424 standard; DNA; 9605 BP.
XX
AC ABK91424;

XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA mutant 1.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; NS5A; NS5B; NS5C; NS5D; NS5E; NS5F;
XX internal ribosome entry site; IRES; NS5A; NS5B; NS5C; NS5D; NS5E;
OS Hepatitis C virus.
OS Synthetic.
XX
PH Key Location/Qualifiers
XX CDS 342..9374
XX /*tag= a
XX /product= "HCV polyprotein"
XX /note= "The polyprotein consists of the Core, E1,
XX E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
XX mutation replace (3625,G)
XX /*tag= b
XX
PN W0200259321-A2.
XX
PD 01-AUG-2002.
XX
PF 16-JAN-2002; 2002WO-EF00526.
XX
PP 23-JAN-2001; 2001US-263479P.
XX
PR (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
PI De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression
XX
PS Claim 9; Page -: 69pp; English.
XX

CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.
XX
SQ Sequence 9605 BP; 1910 A; 2883 C; 2732 G; 2079 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 4 TTGGTAAGGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGTACATCCG 63

Db	696	TTGGGTAAAGGTCAATGCATACCCCTACGTGCGGCTTCGCCGATCTCATGGGGTCAATCCG	755
Qy	64	CTCGTCGGCGCCCTAGGGGGCGCTGCAAGGGCCCTGGGCGCATGCGGTCCGGGTCTG	123
Db	756	CTCGTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTG	815
Qy	124	GAGGACGGCGTGAACATATGCACACAGGAATTTGCGCGGTGCTCTTTCTCTATCTTCCTC	183
Db	816	GAGGACGGCGTGAACATATGCACACAGGAATCTGCCCGGTGCTCTTTCTCTATCTTCCTT	875
Qy	184	TTGGCTTTGCTGTCTGCTGTGACCGTTCAGCTTCGGCTTATGAAGTCGCAACGTGTCC	243
Db	876	TTGGCTTTGCTGTCTGCTTTGACCATCCACAGCTTCGGCTTATGAAGTCGCAACGTATCC	935
Qy	244	GGGATGTACCATGTCTACGAACGACTCTCCAACTCAAGCATTTGTGTATGAGCAGCGAC	303
Db	936	GGAGTGTACCATGTACGAACGACTCTCCAAACGCAAGCATTTGTGTATGAGCAGCGAC	995
Qy	304	ATGATCATGACACACCCCGGGTCGTCCTGGTTTCGGGAGAACAACTCTTTCCCGCTGC	363
Db	996	ATGATCATGACATACCCCGGGTCGTCCTGGTTTCGGGAGAACAACTCTTTCCCGCTGC	1055
Qy	364	TGGGTAGCGTCAACCCCAAGCTCGAGCTAGGAACGCCAGCGTCCCAACACGACATA	423
Db	1056	TGGGTAGCGTCACTCCCAAGCTCGCGCCAGGAACGCTAGCGTCCCACTACGACGATA	1115
Qy	424	CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGGG	483
Db	1116	CGACGCCATGTGATTTGCTCGTTGGGGCGGCTGCTCTCTGCTCCGCTATGTACGTGGGA	1175
Qy	484	GACCTCTGCGGATCTGTCTTCTCGTCCCGCTGTTACACATCTCGCTCGCGGCAT	543
Db	1176	GATCTCTGCGGATCTGTTTCTCGTCGCCACGCTTCACTTCTCGCTCGCGGCAC	1235
Qy	544	GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTGTCGGTATG	603
Db	1236	GAGACGATACAGGACTGCAATTTGCTCAATATATCCCGGCCACGTGACAGGTCAACCGTATG	1295
Qy	604	GCTTCGGGATATGATGATGAACCTGGTCGCTACACAGGCGCTGGTGGTATCGACGCTGCTC	663
Db	1296	GCTTCGGGATATGATGATGAACCTGGTCACTACAGACGCGCTAGTGGTATGCGAGTTACTC	1355
Qy	664	CGGATCCACACAGCTGTCGTGACATGTGTGGCGGGGCCCATATGGGAGATCTTGGCGGGT	723
Db	1356	CGGATCCACACAGCTGTCGTGATATGTGTGGCGGGGCCCATATGGGAGATCTTGGCGGGC	1415
Qy	724	CTCGCCTACTATTCCATGGTGGGGAACGTGGGCTAAGTTTGTATGTGATGCTACTCTTT	783
Db	1416	CTTGCCTACTATTCCATGGTGGGGAACGTGGGCTAAGTTTGTATGTGATGCTACTCTTT	1475
Qy	784	GC 785	
Db	1476	GC 1477	

Query, Acc.	Best Local Matches	Similarity 93.7%	Pred. No. 3.4e-179	Mismatches 0	Indels 49	Gaps 0
RESULT 9	ABK91425	ABK91425 standard; DNA; 9605 BP.				
ID	ABK91425	ABK91425 standard; DNA; 9605 BP.				
XX	XX	XX				
AC	ABK91425;	XX				
XX	XX	XX				
15-NOV-2002	(first entry)	XX				
DT	XX	Hepatitis C virus Con 1 isolate DNA mutant 2.				
DE	XX	Hepatitis C virus				
XX	XX	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;				
KW	KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;				
KW	KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.				
XX	XX	XX				
OS	OS	Hepatitis C virus.				
OS	OS	Synthetic.				
XX	XX	XX				

Db 876 TTGGCTTTGCTGCTCTTTGACATCCAGCTTCCGCTTATGAAGTGCACAGTATCC 935
 Qy 244 GGATGTACCATGTCAGAACGAGTCTCTCAACTCAAGCATTTGTATGAGGACGGAC 303
 Db 936 GGATGTACCATGTCAGAACGAGTCTCTCAAGCATTTGTATGAGGACGGAC 995
 Qy 304 ATGATCATGACACCCCGGGTGGTCCCTGCTTGGGAGAACAACTCTTCCGCTGC 363
 Db 996 ATGATCATGACACCCCGGGTGGTCCCTGCTTGGGAGAACAACTCTTCCGCTGC 1055
 Qy 364 TGGGTAGCGCTCAACCCCGAGCTCGAGCTAGGAACCGCAGCGTCCACACGACAATA 423
 Db 1056 TGGGTAGCGCTCACTCCACGCTCGCGCCAGGAACCTAGCGTCCCACTAGACAGATA 1115
 Qy 424 CGACGCCACGTCGATTTGCTGTTGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGG 483
 Db 1116 CGACGCCACGTCGATTTGCTGTTGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGG 1175
 Qy 484 GACCTCTCGGATCTGCTCTGCTGCTCCAGCTGTTCACCATCTCGCTCGCGGCAT 543
 Db 1176 GATCTCTGCGGATCTGTTTCTGCTGCTCGCCAGCTGTTCACCTTCTCGCTCGCGGCAC 1235
 Qy 544 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGGCCACATACCGGTCAACGATG 603
 Db 1236 GAGACGCTGACGAGCTCAATTTGCTCAATCTATCCCGGCCACGTCAGGTCAACGATG 1295
 Qy 604 GCTTGGGATATGATGAACTGCTGCTCCCTACACCGCCCTGCTGCTGCTGCTGCTGCT 663
 Db 1296 GCTTGGGATATGATGAACTGCTGCTCCCTACACCGCCCTGCTGCTGCTGCTGCTGCT 1355
 Qy 664 CGATCCCAAGCTGCTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
 Db 1356 CGATCCCAAGCTGCTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
 Qy 724 CTGCTCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
 Db 1416 CTGCTCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1475
 Qy 784 GC 785
 Db 1476 GC 1477

RESULT 10

ID ABK91426
 XX ABK91426 standard; DNA; 9605 BP.
 AC ABK91426;
 XX 15-NOV-2002 (first entry)
 DE Hepatitis C virus Con 1 isolate DNA mutant 3.
 XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 XX Hepatitis C virus.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH 342..9374
 FT CDS /*tag= a
 FT /*product= "HCV polyprotein"
 FT /*note= "The polyprotein consists of the Core, E1,
 FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation replace (4380, G)
 FT /*tag= b
 FT WO200259321-A2.
 XX 01-AUG-2002.

XX 16-JAN-2002; 2002WO-EP00526.
 XX 23-JAN-2001; 2001US-263479P.
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication
 PT and expression
 XX Claim 9; Page -; 69pp; English.
 XX The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the
 CC mutations are detailed in the specification. Also included are
 CC (1) an expression vector comprising a nucleotide sequence coding for
 CC the altered nucleic acids, which is transcriptionally coupled to an
 CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
 CC the altered nucleic acids; (3) a recombinant cell produced by introducing
 CC into a human hepatoma cell the altered nucleic acids; (4) producing an
 CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
 CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
 CC method; and (6) measuring the ability of a compound to affect HCV
 CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
 CC studying HCV replication and expression, and HCV and host cell
 CC interactions, producing HCV RNA and proteins, and providing a system
 CC for measuring the ability of a compound to modulate one or more HCV
 CC activities e.g. to discover drugs which may treat HCV mediated
 CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
 CC The present sequence is an HCV replicon Con 1 mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the HCV sequence appearing as
 CC ABK91411 and the information in Claim 9.
 XX Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
 Best Local Similarity 93.7%; Pred. No. 3.4e-179;
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 Qy 4 TTGGGTAAGTGCATCGATACCCCTTACATGCGGCTTCGCCGACCTCTGTTGGGTACATTCCG 63
 Db 696 TTGGGTAAGTGCATCGATACCCCTCAGCTGCGGCTTCGCCGATCTCATGGGTACATTCCG 755
 Qy 64 CTGCTGGCGCCCCCTAGGGGGGCTGCCAGGCCCCCTGGCGCATGCGGCTTCTG 123
 Db 756 CTGCTGGCGCCCCCTAGGGGGGCTGCCAGGCCCCCTGGCGCATGCGGCTTCTG 815
 Qy 124 GAGGACGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTCTC 183
 Db 816 GAGGACGGCTGAACATATGCAACAGGGAATTCGCCCGGTTGCTCTTCTTCTATCTTCT 875
 Qy 184 TTGGCTTTGCTGCTCTGCTCTGACCGTTTCAGCTTTCGCTTTATGAAGTGCACAGTGTCC 243
 Db 876 TTGGCTTTGCTGCTCTGCTTTGACCATCCAGCTTTCGCTTTATGAAGTGCACAGTATCC 935
 Qy 244 GGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGAC 303
 Db 936 GGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGAC 995
 Qy 304 ATGATCATGACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCGCTGC 363
 Db 996 ATGATCATGACACCCCGGGTGGTCCCTGCTGCGGAGAACAACTCTTCCGCTGC 1055
 Qy 364 TGGGTAGCGCTCAACCCCGAGCTCGAGCTAGGAACCGCAGCGTCCACACGACAATA 423

Mon Dec 22 13:28:52 2003

Db 1056 TGGGTAGCGTCACTCCACGCTCGCGGCAGGAACGCTAGCGTCCCACTACGACGATA 1115
Qy 424 CGAGCCACCTCGATTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Db 1116 CGAGCCATGTCGATTGCTCGTTGGGGGGCTGCTCTCTGCTCCGCTATGTACGTGGGA 1175
Qy 484 GACCTCTGCGGATGCTGCTTCCTCGTCTCCAGCTGTTCCACCTGTTCCGCTCGCGGAT 543
Db 1176 GATCTCTGCGGATGCTGTTTCTCTGCTCGCCAGCTGTTCCACCTGTTCCGCTCGCGGAC 1235
Qy 544 GAGACGTCGAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603
Db 1236 GAGACGTCGAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGTATG 1295
Qy 604 GCTTGGGATATGATGAACTGCTCGCTCCGCTACACGCGCCCTGTTGATCGGAGCTGCTC 663
Db 1296 GCTTGGGATATGATGAACTGCTCGCTCCGCTACACGCGCCCTGTTGATCGGAGCTGCTC 1355
Qy 664 CGATCCACAGCTGCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 723
Db 1356 CGATCCACAGCTGCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1415
Qy 724 CTCGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 783
Db 1416 CTTGCTACTATTCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 1475
Qy 784 GC 785
Db 1476 GC 1477

RESULT 11
ABK91428
ID ABK91428 standard; DNA; 9605 BP.
AC ABK91428;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA mutant 5.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX
OS Hepatitis C virus.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FD CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A, NS5B and NS5B proteins"
FT mutation replace (6463.A)
FT /*tag= b
XX
XX WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP00526.
XX PF
XX 23-JAN-2001; 2001US-263479P.
XX PR
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX PA
XX De Francesco R, Migliaccio G, Paonessa G;
XX PI
XX WPI; 2002-599793/64.
XX DR
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV

PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression -
XX
XX Claim 9; Page -: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids; (2) a recombinant cell human hepatoma cell comprising
CC into a human hepatoma cell; the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.
XX
XX Sequence 9605 BP; 1909 A; 2884 C; 2733 G; 2079 T; 0 other;
SQ

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
Qy 4 TTGGGTAAGTTCATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATTCGG 63
Db 696 TTGGGTAAGTTCATCGATACCTTACATGCGGCTTCGCGGACCTCGTGGGTACATTCGG 755
Qy 64 CTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTGGGCGCATGGGCTCGGCTTCG 123
Db 756 CTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTGGGCGCATGGGCTCGGCTTCG 815
Qy 124 GAGGAGCGGTGAACTATGCAACAGGAAATTCGCGGTTGCTCTTCTATCTATCTTCCT 183
Db 816 GAGGAGCGGTGAACTATGCAACAGGAAATTCGCGGTTGCTCTTCTATCTATCTTCCT 875
Qy 184 TTGGCTTTGCTGCTCTGCTGTGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGTGTC 243
Db 876 TTGGCTTTGCTGCTCTGCTGTGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGTGTC 935
Qy 244 GGGATGTACCATGTACGAAACGATGCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 303
Db 936 GGAGTGTACCATGTACGAAACGATGCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 995
Qy 304 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 996 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
Qy 364 TGGGTAGCGCTCACCCCGACGCTCGAGCTAGGACGCCAGCGTCCCGACACGACATA 423
Db 1056 TGGGTAGCGCTCACCCCGACGCTCGAGCTAGGACGCCAGCGTCCCGACACGACATA 1115
Qy 424 CGAGCCACCTCGATTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Db 1116 CGAGCCATGTCGATTGCTCGTTGGGGGGCTGCTCTCTGCTCCGCTATGTACGTGGGA 1175
Qy 484 GACCTCTGCGGATGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db 1176 GATCTCTGCGGATGCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1235
Qy 544 GAGACGTCGAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603

Db 1236 GAGACAGTACAGGACTGCAATTGCTCAATATATCCCGGCCACGTGACAGGTACCCGATG 1295
QY 604 GCTTGGGATATGATGAACTGGTCCCTACAAACGGCCCTGGTGGTATCCGAGCTGCTC 663
Db 1296 GCTTGGGATATGATGAACTGGTCCCTACAAACGGCCCTGGTGGTATCCGAGTACTC 1355
QY 664 CGATCCCAACAGCTGCTGGGATGATGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 723
Db 1356 CGATCCCAACAGCTGCTGGGATGATGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 1415
QY 724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAAGGTTTGGATGGTACTCTTT 783
Db 1416 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAAGGTTTGGATGGTACTCTTT 1475
QY 784 GC 785
Db 1476 GC 1477

RESULT 12
ABK91429
ID ABK91429 standard; DNA; 9605 BP.
XX
AC ABK91429;
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA mutant 6.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT FT /*tag= a
FT FT /product= "HCV polyprotein"
FT FT /note= "The polyprotein consists of the Core, E1,
FT FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT FT mutation replace (6859,C)
FT FT /*tag= b
XX
XX WO200259321-A2.
XX
PD 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP00526.
XX
XX 23-JAN-2001; 2001US-263479P.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication
XX and expression
XX
XX Claim 9; Page -: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the
XX mutations are detailed in the specification. Also included are
XX (1) an expression vector comprising a nucleotide sequence coding for

CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC Note: The present sequence is an HCV replicon Con 1 mutant of the invention.
CC The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.
XX
SQ Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 other;
Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 4 TTGGGTAAAGTCAATGATACCTTACATGGCGGCTTGCAGCAGCTCGTGGGTACATCCG 63
Db 696 TTGGGTAAAGTCAATGATACCTTACATGGCGGCTTGCAGCAGCTCGTGGGTACATCCG 755
QY 64 CTGCTGGCGGCCCTAGGGGGCGCTGCCAGAGCCCTGGCGCATGGCGTTCG 123
Db 756 CTGCTGGCGGCCCTAGGGGGCGCTGCCAGAGCCCTGGCGCATGGCGTTCG 815
QY 124 GAGGAGCGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTCTC 183
Db 816 GAGGAGCGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTCTC 875
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGGTTTCAGCTTCGCTTATGAGTGGCGCAACGTGTC 243
Db 876 TTGGCTTTGCTGCTGCTGCTGACCGGTTTCAGCTTCGCTTATGAGTGGCGCAACGTATCC 935
QY 244 GGGATGTACCATGCTCAGAACGAGTCTCCAACTCAAGCAATTTGATAGGAGCGGAC 303
Db 936 GGGATGTACCATGCTCAGAACGAGTCTCCAACTCAAGCAATTTGATAGGAGCGGAC 995
QY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 996 ATGATCATGACATACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
QY 364 TGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAAGCCAGCGTCCCGACAGACAATA 423
Db 1056 TGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAAGCCAGCGTCCCGACAGACAATA 1115
QY 424 CGACGCCAGCTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 1116 CGACGCCAGCTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
QY 484 GACCTCTGCGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db 1176 GATCTCTGCGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1235
QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATG 603
Db 1236 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACCGTACAGGTATG 1295
QY 604 GCTTGGGATATGATGAACTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 663
Db 1296 GCTTGGGATATGATGAACTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1355
QY 664 CGATCCCAACAGCTGCTGGGATGATGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723
Db 1356 CGATCCCAACAGCTGCTGGGATGATGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1415
QY 724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAAGTTTTGGATGGTACTCTTT 783

Mon Dec 22 13:28:52 2003

Db 1416 CTTGGCTACTATTCAGTGGGGAACGGCTAGGTTCTGATGTGATGCTACTCTTT 1475

QY 784 GC 785
Db 1476 GC 1477

RESULT 13
ID ABK91430 standard; DNA; 9605 BP.

XX AC ABK91430;

XX 15-NOV-2002 (first entry)

XX Hepatitis C virus Con 1 isolate DNA mutant 7.

XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

XX Hepatitis C virus.
OS Synthetic.

XX Key Location/Qualifiers
FT CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT mutation replace (6911,C)
FT /*tag= b

XX WO200259321-A2.

XX 01-AUG-2002.

XX 16-JAN-2002; 2002WO-EP00526.

XX 23-JAN-2001; 2001US-263479P.

XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.

XX De Francesco R, Migliaccio G, Paonessa G;

XX WPI; 2002-599793/64.

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression -

XX Claim 9; Page -; 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS4A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated

CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.

XX Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 other;

Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGTTAAGTTCATCGATACCCCTTATCATCGGCTTCGCCGACCTCGTGGGGTACATTCGG 63

Db 696 TTGGTTAAGTTCATCGATACCCCTTATCATCGGCTTCGCCGACCTCGTGGGGTACATTCGG 755

QY 64 CTCGTCGGCGCCCCCTTAGGGGCGCTGCGCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123

Db 756 CTCGTCGGCGCCCCCTTAGGGGCGCTGCGCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 815

QY 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTCCCGGTTGCTCTTCTATCTATCTCTTC 183

Db 816 GAGGACGGCGTGAACATATGCAACAGGGAATTTCCCGGTTGCTCTTCTATCTATCTCTTC 875

QY 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAAGCTGTC 243

Db 876 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAAGCTGTC 935

QY 244 GGGATGTACCATGTACGAAAGGACTGCTCAAGCATTTGATGATAGGAGGAGGAGGAC 303

Db 936 GGGATGTACCATGTACGAAAGGACTGCTCAAGCATTTGATGATAGGAGGAGGAGGAC 995

QY 304 ATGATCATGCACACCCCGGGTGGTGGCTTCCGAGTTCGGGAGAACAACTCTTCCCGTGC 363

Db 996 ATGATCATGCATACCCCGGGTGGTGGCTTCCGAGTTCGGGAGAACAACTCTTCCCGTGC 1055

QY 364 TGGGTAGCGCTCACCCCGACGCTGCGAGTAGGAACGCCAGCGTCCACCAACACACAATA 423

Db 1056 TGGGTAGCGCTCACCCCGACGCTGCGAGTAGGAACGCCAGCGTCCACCAACACACAATA 1115

QY 424 CGACGCCAGTTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483

Db 1116 CGACGCCAGTTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 1175

QY 484 GACCTCTGCGGATGTCTTCTGCTGCTGCGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 543

Db 1176 GATCTCTGCGGATGTCTTCTGCTGCTGCGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 1235

QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATTAACCGGTCACCGTATG 603

Db 1236 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATTAACCGGTCACCGTATG 1295

QY 604 GCTTGGGATATGATGATGAACCTGGTGCCTCAACGGCCCTGGTGGTATCGCAGCTGTC 663

Db 1296 GCTTGGGATATGATGATGAACCTGGTGCCTCAACGGCCCTGGTGGTATCGCAGCTGTC 1355

QY 664 CGGATCCCAAGCTGCTGGGACATGTTGGGGGGGCCCATTTGGGGAGTCTCTGGCGGGT 723

Db 1356 CGGATCCCAAGCTGCTGGGACATGTTGGGGGGGCCCATTTGGGGAGTCTCTGGCGGGT 1415

QY 724 CTGCGCTACTATTCCATGTTGGGGAATCGGGCTTAAGGTTTTGATTTGATGCTACTCTTT 783

Db 1416 CTGCGCTACTATTCCATGTTGGGGAATCGGGCTTAAGGTTTTGATTTGATGCTACTCTTT 1475

QY 784 GC 785

Db 1476 GC 1477

RESULT 14
ABK91431
ID ABK91431 standard; DNA; 9605 BP.
XX

Mon Dec 22 13:28:52 2003

XX PH Location/Qualifiers
FT CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /notes= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT mutation replace (6936,G)
FT /*tag= b
XX W0200259321-A2.
XX 01-AUG-2002.
XX 16-JAN-2002; 2002WO-EP00526.
XX 23-JAN-2001; 2001US-263479P.
XX (RICE-) IST RIBERCHIE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication
XX and expression -
XX Claim 9; Page -; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS4A, or EMCV IRES mutations, respectively. The location of the
XX mutations are detailed in the specification. Also included are
XX (1) an expression vector comprising a nucleotide sequence coding for
XX the altered nucleic acids, which is transcriptionally coupled to an
XX exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
XX the altered nucleic acids; (3) a recombinant cell produced by introducing
XX into a human hepatoma cell the altered nucleic acids; (4) producing an
XX HCV (hepatitis C virus) replicon enhanced cell or which containing a
XX functional HCV replicon; (5) an HCV replicon enhanced cells made in the
XX method; and (6) measuring the ability of a compound to affect HCV
XX activity. The HCV replicons and HCV replicon enhanced cells are useful in
XX studying HCV replication and expression, and HCV and host cell
XX interactions, producing HCV RNA and proteins, and providing a system
XX for measuring the ability of a compound to modulate one or more HCV
XX activities e.g. to discover drugs which may treat HCV mediated
XX diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
XX The present sequence is an HCV replicon Con 1 mutant of the invention.
XX Note: The present sequence is not shown in the specification but
XX was created by the indexer using the HCV sequence appearing as
XX ABK91411 and the information in Claim 9.
XX SQ Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 other;
Query Match 88.5%; Score 703.6; DB 24; Length 9605;
Best Local Similarity 93.7%; Pred. No. 3.4e-179; Indels 0; Gaps 0;
Matches 733; Conservative 0; Mismatches 49;
4 TTGGGTAAAGTTCATGCATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATCCG 63
696 TTGGGTAAAGTTCATGCATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATCCG 755
64 CTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGTCCGGGTTCTG 123
756 CTCGTCGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGTCCGGGTTCTG 815
124 GAGGACGGCGTGAACATATGCAACAGGGAATTCGCGGGTTCCTTTCTATCTCCTC 183
816 GAGGACGGCGTGAACATATGCAACAGGGAATTCGCGGGTTCCTTTCTATCTCCTT 875

QY 184 TTGGCTTTGCTGTCCTGCTGACCGTTCAGCTTCCGCTTATGAAGTCGCGCAACGTGTCC 243
Db |||||
QY 876 TTGGCTTTGCTGTCCTGCTGACCGTTCAGCTTCCGCTTATGAAGTCGCGCAACGTATCC 935
Db |||||
QY 244 GGGATGTACATGTCAAGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 303
Db |||||
QY 936 GAGGTGTACATGTCAAGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 995
Db |||||
QY 304 ATGATCATGCACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
Db |||||
QY 996 ATGATCATGCACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGCTGC 1055
Db |||||
QY 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGTCCCGACACGACAATA 423
Db |||||
QY 1056 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGTCCCGACACGACAATA 1115
Db |||||
QY 424 CGACGCCACGTCGATTTGCTGCTGGGGGGGCTGCTTCTGCTTCCGCTATGTACGTTGGG 483
Db |||||
QY 1116 CGACGCCACGTCGATTTGCTGCTGGGGGGGCTGCTTCTGCTTCCGCTATGTACGTTGGG 1175
Db |||||
QY 484 GACCTCTGCGGATCTGCTTCTCTCCGCTGCTCCAGCTGTTCCACATCTCGCTCGCGGCA 543
Db |||||
QY 1176 GATCTCTGCGGATCTGCTTCTCTCCGCTGCTCCAGCTGTTCCACATCTCGCTCGCGGCA 1235
Db |||||
QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATACGGGTCAACGTATG 603
Db |||||
QY 1236 GAGACAGTACAGGACTGCAATTTGCTCAATATATCCCGGCGACATACGGGTCAACGTATG 1295
Db |||||
QY 604 GCTTGGGATATGATGAACTGCTGCGCTACAAACGCGCCCTGGTGGTATCGCAGCTGCTC 663
Db |||||
QY 1296 GCTTGGGATATGATGAACTGCTGCGCTACAAACGCGCCCTGGTGGTATCGCAGTTACTC 1355
Db |||||
QY 664 CGGATCCCAAGCTGCTGCGGAGCATGCTGGCGGGGCGCCATTTGGGAGTCTCGGGGGT 723
Db |||||
QY 1356 CGGATCCCAAGCTGCTGCGGAGCATGCTGGCGGGGCGCCATTTGGGAGTCTCGGGGGC 1415
Db |||||
QY 724 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAAGCTTTTGAATTTGTATGCTACTCTTT 783
Db |||||
QY 1416 CTTGCTACTATTCCATGCTGGGAACTGGGCTAAAGCTTTCTGATTTGTATGCTACTCTTT 1475
Db |||||
QY 784 GC 785
Db |||||
QY 1476 GC 1477
Db |||||

Search completed: December 19, 2003, 18:51:03
Job time : 222.785 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 2045.18 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-5
Perfect score: 795
Sequence: 1 ATGTGGGTAAAGTATCGA.....TACTCTTTCCTCCTTAATAG 795

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
C 1	115.6	14.5	488	9	AV755731
C 2	91.6	11.5	492	9	AV758366
C 3	47.2	5.9	1201	13	BX356664
C 4	43.4	5.5	925	29	CNS0091P

5	42.8	5.4	423	9	AA459034
C 6	41.6	5.2	1201	9	AL513886
7	41.2	5.2	1195	28	BI0902
8	40.6	5.1	402	9	AV392783
9	40.6	5.1	551	9	AV392165
10	40.6	5.1	552	12	BI996341
11	40.6	5.1	584	12	BI727879
12	40.6	5.1	525	10	BE337089
13	40.6	5.0	534	14	CD040840
14	40.6	5.0	671	12	BI723733
C 15	39.8	5.0	608	14	CB640103
C 16	39.8	5.0	738	14	CB668031
C 17	39.8	5.0	969	12	BM017656
C 18	39.6	5.0	675	12	BQ45314
19	39.2	4.9	571	12	BM692316
20	39.2	4.9	645	29	CNS01213
21	39.2	4.9	1102	13	BQ069967
C 22	39	4.9	359	12	BJ252669
23	39	4.9	375	12	BJ246716
24	39	4.9	621	14	CA816001
25	39	4.9	624	12	BI723734
26	39	4.9	840	29	CC335916
27	39	4.9	873	14	CD446071
28	38.8	4.9	590	14	CA659369
29	38.4	4.8	702	14	CD432549
C 30	38.4	4.8	740	12	BJ536071
31	38.4	4.8	970	29	CNS010C9
C 32	38.4	4.8	987	29	CNS015VX
C 33	38.2	4.8	1270	12	BG968359
C 34	38	4.8	354	14	CB966525
C 35	38	4.8	742	13	BQ752673
36	37.8	4.8	435	14	C72860
37	37.8	4.8	533	29	CC010084
38	37.8	4.8	659	29	CC405164
39	37.8	4.8	826	29	BZ736582
C 40	37.8	4.8	895	29	CC359028
41	37.8	4.8	925	29	CC359026
C 42	37.8	4.8	940	29	CC010085
C 43	37.8	4.8	951	29	CC405167
44	37.8	4.8	1105	13	BQ064949
45	37.8	4.8	1126	13	BQ063359

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAK903 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

source	1. .488	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="BMFAK03"	
		/tissue_type="Bone marrow"	
		/cell_type="CD34+ hematopoietic stem/progenitor cell"	
		/lab_host="BM25.8"	
		/clone_lib="BM"	
		/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"	
BASE COUNT	116 a 134 c 137 g 97 t		
ORIGIN			
Query Match	14.5%; Score 115.6; DB 9; Length 488;		
Best Local Similarity	64.5%; Pred. No. 6.4e-19; Indels 4; Gaps 2;		
Matches	205; Conservative		
445	GTGGGGGGCTGCTTTCTGTTCCGCTATGACGTGGGGACCTCTGCGATCTGCTTC		504
472	GTGGGTACACTCGCTCTGCTCAGCTCTCTACGTGTGGGACCTCTGCGAGGATGATG		413
505	CTCGTCTCCAGCTGTTACCACTCTCGCTCGCCGGCATGAGCGTGCAGACTGCAAT		564
412	CTTGCAAGTTCAAGCTG---ATCATCTGGGCTCAGCACCATGAGTTTGTGATGAATGCAC		356
565	TGCTCAATCTATCCGGCCACATAACGGGTACCCGCTATG-GCTGGGATATGATGAA		623
355	TGCTCAATCTATCTGGCCCATCACTGACACCGTATGAGCATGGGACATGATGAA		296
624	CTGGTCGCTACAAAGCCCTGTGTGATCGAGCTGCTCCGATGCCACAAAGCTGTCT		683
295	CTGGTCGTGACCGCTGCTATGATCATGATGCGTACGCAATGCGGTTCTCTGAGTATCAT		236
684	GGATCTGGTGGGGGGCCCATTTGGGGAGTCTTGGGGGCTCTGCGCTACTATTCACATGT		743
235	AGATATATACAGCGGGGCTCACTGGGGGCTCATGTTCCGCTTAGCCTACTTCTCTATGCA		176
744	GGGGAATCGGGCTAAGGT 761		
175	GGGAGCTGGCGCAAGT 158		
RESULT 2			
LOCUS	AV758366/c		
DEFINITION	AV758366 BM Homo sapiens CDNA clone BMFAK03 5', mRNA sequence.		
ACCESSION	AV758366		
VERSION	AV758366.1 GI:10916214		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1. (bases 1 to 492)		
JOURNAL	Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,		
COMMENT	Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng		
	Gu, Y., Li, N., Qian, B., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,		
	L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,		
	Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.		
	Homo sapiens CDNA BM clones		
	Unpublished		
	Contact: Zeguang Han		
	Chinese National Human Genome Center at Shanghai		
	351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai		
	201203 P. R. China		
	Tel: 86-21-50801919 (ex. 45)		
	Fax: 86-21-50801922		
	Email: hanzg@hgc.sh.cn		
	This clone is available at CHGC in Shanghai.		
	Location/Qualifiers		
FEATURES	1. .492		
source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="BMFAK03"		
	/tissue_type="Bone marrow"		
	/cell_type="CD34+ hematopoietic stem/progenitor cell"		
	/lab_host="BM25.8"		
	/clone_lib="BM"		
	/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"		
BASE COUNT	116 a 134 c 137 g 97 t		
ORIGIN			
Query Match	14.5%; Score 115.6; DB 9; Length 488;		
Best Local Similarity	64.5%; Pred. No. 6.4e-19; Indels 4; Gaps 2;		
Matches	205; Conservative		
445	GTGGGGGGCTGCTTTCTGTTCCGCTATGACGTGGGGACCTCTGCGATCTGCTTC		504
472	GTGGGTACACTCGCTCTGCTCAGCTCTCTACGTGTGGGACCTCTGCGAGGATGATG		413
505	CTCGTCTCCAGCTGTTACCACTCTCGCTCGCCGGCATGAGCGTGCAGACTGCAAT		564
412	CTTGCAAGTTCAAGCTG---ATCATCTGGGCTCAGCACCATGAGTTTGTGATGAATGCAC		356
565	TGCTCAATCTATCCGGCCACATAACGGGTACCCGCTATG-GCTGGGATATGATGAA		623
355	TGCTCAATCTATCTGGCCCATCACTGACACCGTATGAGCATGGGACATGATGAA		296
624	CTGGTCGCTACAAAGCCCTGTGTGATCGAGCTGCTCCGATGCCACAAAGCTGTCT		683
295	CTGGTCGTGACCGCTGCTATGATCATGATGCGTACGCAATGCGGTTCTCTGAGTATCAT		236
684	GGATCTGGTGGGGGGCCCATTTGGGGAGTCTTGGGGGCTCTGCGCTACTATTCACATGT		743
235	AGATATATACAGCGGGGCTCACTGGGGGCTCATGTTCCGCTTAGCCTACTTCTCTATGCA		176
744	GGGGAATCGGGCTAAGGT 761		
175	GGGAGCTGGCGCAAGT 158		
RESULT 3			
LOCUS	BX356664		
DEFINITION	BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA		
ACCESSION	BX356664		
VERSION	BX356664.1 GI:30378083		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1. (bases 1 to 1201)		
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by life Technologies, a division of		
	Invitrogen. Contact : Feng liang Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
	Paradise Avenue Genoscope sequence ID : CSODI015CA02NP1.		
	Location/Qualifiers		

melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

COMMENT

; eurosides II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1195)
REFERENCE Peng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
AUTHORS Ecker, J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished
COMMENT Other_GSSs: F13A13-T7
Contact: Ecker, J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@genome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 76
High quality sequence stop: 98.
Location/Qualifiers
1. 1195
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F13A13"
/sex="hermaphrodite"
/clone_lib="ICP"
/notes="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT 61 a 541 c 178 g 404 t 11 others
ORIGIN
Query Match 5.2%; Score 41.2; DB 28; Length 1195;
Best Local Similarity 50.0%; Pred. No. 8.2;
Matches 103; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 332 CTGGCGTTGGGAGAACACTCTTCCCGTCTGGGTAGCGCTCACCCACGCTCCGAG 391
DB 295 CTTTCTTTTCGGTGTACGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 354
QY 392 CTAGGACGCGAGCTCCACACGACATACAGCCAGCTCTCGGAGCTCTCTCTCTCTCT 451
DB 355 CTTGGGCTCTCCGCCCCCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 414
QY 452 CGGCTGCTTTCTGTCGCTATGACGTGGGGACCTCTCGGAGCTCTCTCTCTCTCTCTCT 511
DB 415 CCGTTCCTTTCTGCGCGCTCTCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 474
QY 512 CCAGCTGTTTACCATCTCGCTGCG 537
DB 475 CTTTCTTTTTCGCGCGCGCTGCG 500
RESULT 8
AV392783
LOCUS AV392783 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION CDNA clone CM086504_r 5', mRNA sequence.
ACCESSION AV392783
VERSION AV392783.1 GI:6546999
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Rukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 402)
REFERENCE Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
AUTHORS A large scale structural analysis of cDNAs in a unicellular green
TITLE alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
JOURNAL DNA Res. 6 (6), 369-373 (1999)
MEDLINE 20152988
PUBMED 10691129
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

10691129
PUBMED
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
1. 402
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM086504_r"
/dev_stage="photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 66 a 126 c 146 g 64 t
ORIGIN
Query Match 5.1%; Score 40.6; DB 9; Length 402;
Best Local Similarity 45.3%; Pred. No. 8.2;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 42 CGACCTGCTGGGTACATTCCTGCTGCGCGCCCTAGGGGGCGCTCCAGGGCCCT 101
DB 53 CGAGCTCATCTCTGTCATTTGTGCGCGCACTGCCAACATGAGGACGCTGCTGACGGACCT 112
QY 102 GCGCATGGCTCCGGTCTCTGGAGGACGCGCTGAATATGCAACAGGGAATTTGCCCGG 161
DB 113 GCGCGGCGCGCGAGTGGAGGCGCGCTACGGCGACGAGTCCCTGAGCTTGGCGCG 172
QY 162 TTGCTCTTTCTCTATCTCTCTTGGCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 221
DB 173 CCGCAAGGTCTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCAAGACCCAG 232
QY 222 TTATGAAGTCGCGCAACGTGTCGGGATGTACCATGTACCAAGAGCTGCTTCAACTCAAG 281
DB 233 CTTTCGCGCTCCGCTGCTGCGCCACTGCTGCGCGCGCGCGCGCGCTGCTGCTGCTGCT 292
QY 282 CATTGTGTATGAGCGCGGACATGATCATGACACCCCGGTGGTGGCTGCGCTGCGCTGCG 341
DB 293 CCTGATGCACACGAGGAGTTCGCGCGCGATCTACGGCGCGGTGCTGCGCGCGCGCGG 352
QY 342 GGAGAACAACTCTTCCCGCTGCTGGGT 368
DB 353 CAGAGACGACGAGGACGCTACATGAT 379
RESULT 9
AV392165
LOCUS AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION CDNA clone CM08605_r 5', mRNA sequence.
ACCESSION AV392165
VERSION AV392165.1 GI:6546381
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Rukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 551)
REFERENCE Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
AUTHORS A large scale structural analysis of cDNAs in a unicellular green
TITLE alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
JOURNAL DNA Res. 6 (6), 369-373 (1999)
MEDLINE 20152988
PUBMED 10691129
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Mon Dec 22 13:28:53 2003

Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source

Location/Qualifiers
1. .551
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM083e05_r"
/dev_stage="photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 94 a 182 c 189 g 85 t 1 others
ORIGIN

Query Match 5.1%; Score 40.6; DB 9; Length 551;
Best Local Similarity 45.3%; Pred. No. 9.1;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCTGGGGTACATTCGCTCGTGGCGGCCCTAGGGGGGGCTGCCAGGGCCCT 101
DB 108 CGAGCTCATCTGTCATTTGTGGCGGCACTGCCAATGAAGACGTGCTGACGGACCT 167
QY 102 GCGCATCGCGTCCGGGTTCTGGAGGACGGCGTGAATATGCAACAGGGAATTTGCCCGG 161
DB 168 GCGCGCGCGCGCGAGTGGAGGGCGGCTACGCGCAGAGTCCGTGAGCTTGGCGCG 227
QY 162 TTGCTCTTTCTATCTTCTTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 221
DB 228 CCGCAAGGTGTTTGAAGAGATCAAGAGTACGTGTGAACTCAAGCCCAAGAACCCAG 287
QY 222 TTATGAAGTGGCAACGTTGTCGGGATGATACCATGTACGAAACGACTGCTCCAACTCAAG 281
DB 288 CTTCGCGCTCGCTGGTGGGCACTCGCTGGCGCGGCGCACTGAGTGGCTGCTGCTGAT 347
QY 282 CATTTGTATGAGGACGCGGACATGATCATGACACACCCCGGTCGTCGCTGCTGCTG 341
DB 348 CCTGATGCCACGACGAGAGTTTGGCGCGGCGCATCTACGCGCGCTGCTCCATGCGCGG 407
QY 342 GGAGAACAACTCTTCCCGCTGCTGGGT 368
DB 408 CAAGAGACGACGAGGAGCTACATGAT 434

RESULT 10
BI996341 552 bp mRNA linear EST 25-OCT-2001
LOCUS 1031037A07.Y2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI996341
VERSION BI996341.1 GI:16431115
KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

REFERENCE 1 (bases 1 to 552)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
UNICELLULAR SYSTEM FOR ANALYZING GENE FUNCTION AND REGULATION IN
VASCULAR PLANTS. Project: 1031

JOURNAL Unpublished
COMMENT Contact: Charles Hauser
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES
source 1. .552
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) plant
phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
polyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

BASE COUNT 93 a 184 c 189 g 86 t
ORIGIN

Query Match 5.1%; Score 40.6; DB 12; Length 552;
Best Local Similarity 45.3%; Pred. No. 9.1;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCTGGGGTACATTCGCTCGTGGCGGCCCTAGGGGGGGCTGCCAGGGCCCT 101
DB 110 CGAGCTCATCTTGTGTCATTTGTGGCGGCACTGCCAATGAAGACGTGCTGACGGACCT 169
QY 102 GCGCATCGCGTCCGGGTTCTGGAGGACGGCGTGAATATGCAACAGGGAATTTGCCCGG 161
DB 170 GCGCGCGCGCGCGAGTGGAGGGCGGCTACGCGCACGAGTCCGTGAGCTTGGCGGC 229
QY 162 TTGCTCTTCTATCTTCTTCTTGGCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 221
DB 230 CCGCAAGGTGTTTGAAGAGATCAAGAGTACGTGTGAACCTCAAGGCGCAGAACCCAG 289
QY 222 TTATGAAGTGGCAACGTTGTCGGGATGTACCATGTACGACGACTGCTCCAACTCAAG 281
DB 290 CTTCGCGCTCGCTGGTGGGCACTCGCTGGCGCGGCGCATCTACGCGCGGTGCCCATGCCGG 349
QY 282 CATTTGTATGAGGACGCGGACATGATCATGACACACCCCGGTCGTCGCTGCTGCTGCTG 341
DB 350 CCTGATGCCACGACGAGGAGTTTGGCGCGGCGCATCTACGCGCGGTGCCCATGCCGG 409
QY 342 GGAGAACAACTCTTCCCGCTGCTGGGT 368
DB 410 CAAGAGACGACGAGGAGCTACATGAT 436

RESULT 11
BI727879 584 bp mRNA linear EST 19-SEP-2001
LOCUS 1031095C12.y1 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BI727879
VERSION BI727879.1 GI:15703574
KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

REFERENCE 1 (bases 1 to 584)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
UNICELLULAR SYSTEM FOR ANALYZING GENE FUNCTION AND REGULATION IN
VASCULAR PLANTS. Project: 1031

JOURNAL Unpublished
COMMENT Contact: Charles Hauser

DCWB Box 91000
 Duke University
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu

FEATURES
 source

Location/Qualifiers
 1..584
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mrna"
 /strain="CC-1690 wild type mt+ 2lgr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
 /note="vector: pbluescript II SK-; Site 1: EcoRI; Site 2:
 XhoI; Stress condition II library, constructed by John
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690
 cells grown to mid-log phase in TAP (NH4+ - containing)
 and shifted to TAP - NO3- (24hrs); H2 production
 conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
 PolyA mRNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
 sites. pBluescript II SK- plasmids were excised from the
 lambda Zap clones by superinfection with ExAssist
 (Stratagene) phage. The library was normalized using
 method 4 described in Bonaldo et al., (1996) Genome
 Research 6: 791-806."

BASE COUNT 106 a 188 c 197 g 93 t
 ORIGIN

Query Match 5.1%; Score 40.6; DB 12; Length 584;
 Best Local Similarity 45.3%; Pred. No. 9.3;
 Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTACATTCGCTCGTCGGCGCCCTTAGGGGCGCTGCCAGGCGCCT 101
 Db |||||
 47 CGAGCTCATCTGTGTCATTTGCGCGGCACCTGCCAACATGAAGACCGTCTGCAGCGACCT 106
 QY |||||
 102 GGCGCATGGCGTCCGGTTCCTGGAGGACGGGTGAATCATGTCAACAGGGGAATTTGCCCGG 161
 Db |||||
 107 GCGCGGCGCGCGCGAGTGGAGGGCGGTACGCCACAGTCTGTAGCTTCCAGCTTGGCGC 166
 QY |||||
 162 TTGCTCTTTCTTAICTTCCTTTGGCTTTGCTGTCTGTCTGACCGTTCAGCTTCCGC 221
 Db |||||
 167 CCGCAAGGTGTTTACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCAAGAACCCAG 226
 QY |||||
 222 TTATGAAGTGGCAACGTGTCCGGGATGTACCATGTACGAACGACTGTCTCAACTCAAG 281
 Db |||||
 227 CTTGCGCGTCCGCTGCGTGGGCCACTCGCTGGGCGGGCACCGCCGCGCTGCCTGTCCAT 286
 QY |||||
 282 CATTTGTGTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGCGCTCGGTTTCG 341
 Db |||||
 287 CCTGATGCACCAACAGCAGGAGTTTGGCGCGGCATCTACGCGCGCGTGCCTCATGCCGG 346
 QY |||||
 342 GGAGAACAACTTTTCCGCTGCTGGT 368
 Db |||||
 347 CAAGAAGCAAGGGCAGCTACATGAT 373

RESULT 12	BE337089	525 bp	mRNA	linear	EST 14-JUL-2000
LOCUS	BE337089				
DEFINITION	894043G08.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II				
	Chlamydomonas reinhardtii cDNA, mRNA sequence.				
ACCESSION	BE337089				
VERSION	BE337089.1				
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii				
ORGANISM	Chlamydomonas reinhardtii				
	Eukaryota; Viridiplantae;				
	Chlorophyta; Chlorophyceae; Volvocales;				

Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 525)

Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P., McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.

Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 2

Unpublished

Contact: Elizabeth H. Harris

DCMB Box 91000

Duke University

Durham, NC 27708-1000, USA

Tel: 919 613 8164

Fax: 919 613 8177

Email: chlamy@duke.edu.

Location/Qualifiers

1. 525

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"

II"

/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with EXAssist (Stratagene) phase. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

```

BASE COUNT      99 a   181 c   160 g   85 t
ORIGIN
Query Match      5.0%; Score 40; DB 10; Length 525;
Best Local Similarity 47.0%; Pred. No. 13;
Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 261 GAACGACTGCTCCAACCTCAAGCACTTGTGTATGAGCGAGCGACATGATCATGCACACCCC 320
Db 43 GCACCGCTTCAACAACCCACCCTGCTGANTTCGCCGCCCTACAAGTACCTGCTGTCCAT 102
Qy 321 CGGGTGGTGCCCTGCGTTCCGGAGAAACAACACTCTTCCCGCTGCTGGGTAGCGGTCAACCCC 380
Db 103 GGAGGACAAGACCTGCCACCCTGGAGGTCATCGAGTACTGCGAGCTGGGCAACCTGTCCAA 162
Qy 381 CAGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACACATACGACGCCACCTCGATTT 440
Db 163 CGCGCTCAAAAAACAACATCTTCATGATTCCTCAACCCCGTCAATCGCGCGCGCGCGCGC 222
Qy 441 GCTCGTTGGGGCGGCTGCTTTCTGTTCGGCTAATGTACGTGGGGACCTCTGCGGATCTGT 500
Db 223 GGGCGACGGCGCGCGCGCAGCGTAGCGGAGCGCGCGCAGAGCCCATGAGGT 282
Qy 501 CTTCTCTGCTTCCAGCTGTTTCAAC 524
Db 283 CAACATGCGCACCGCTGCTGCTCAC 306

RESULT 13
CD040840
LOCUS
DEFINITION
pSHB036XB09f 300663 psHB: Infected hypocotyl soybean host. 48 hrs
post infection Phytophthora sojae cDNA clone SHB036B09 5, mRNA
sequence.
ACCESSION
CD040840
VERSION
CD040840.1 GI:30502701
KEYWORDS
EST.

```

Mon Dec 22 13:28:53 2003

Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 671)

Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1031

Unpublished

Contact: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

1. .671

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

Research 6: 791-806."

99 t

BASE COUNT 139 a 224 c 209 g 50 t

ORIGIN

Query Match 5.0%; Score 40; DB 12; Length 671;

Best Local Similarity 47.0%; Pred. No. 14;

Matches 124; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 261 GARACGACTGCTCAACTCAAGCATTTGTATGAGGACGGGACATGATCATGCACACCCC 320

Db 353 GCACCGCTTCAACACACCCACCGTGTGAATTCGCCCGCTTACAGTACTGCTGCTCCAT 412

Qy 321 CGGGTGTGCGCTGCTGCTGGGAGAACAACTTTCCCGCTGCTGGGTAGCGCTCACCCC 380

Db 413 GGAGGACAGACCTTGCACCTGGAGTCTGAGTACTGCGACCTGGGCAACCTGTCCAA 472

Qy 381 CACGCTCGAGCTAGGAAGCCAGCGTCCACACAGACATACGACGACCGACGTCGATTT 440

Db 473 CGGCTCAAAAACAAACATCTTCATGATCCCAACCCCGCTCATCGCGCGCGCGCGCGC 532

Qy 441 GCTGTTGGGCGGCTGCTTCTTCTCCGCTATGTACGTGGGGGACCTCTGGGATCTGT 500

Db 533 GGGGACGCGCGCGCGCGGAGTAGGAGCGCGCGCGCGCGCGAGCAGCCATGAAGT 592

Qy 501 CTTCCTGCTCTCCAGCTGTTCAC 524

Db 593 CAACATGGCACCTGCTGCTAC 616

RESULT 15

CB640103/c

LOCUS

DEFINITION

OSJNEal3E06.f OSJNEa Oryza sativa (japonica cultivar-group) cDNA clone OSJNEal3E06 5', mRNA sequence.

ACCESSION

CB640103

Phytophthora sojae

Phytophthora sojae

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; phytophthora.

1 (bases 1 to 534)

Tyler, B.M., Judelson, H.S., Gijzen, M., Dean, R.A. and Waugh, M.E.

USDA-IFAPs: Expression of Phytophthora sojae genes during infection and propagation

Unpublished

Contact: Tyler B

Tyler lab

VRI

1880 Pratt Dr., Blacksburg, VA 24061, USA

Tel: 540-231-7318

Email: bmtyer@vt.edu

PCR PRIMERS

FORWARD: BK reverse

Plate: 036 row: B column: 09

Seq primer: BK reverse

High quality sequence stop: 534.

Location/Qualifiers

1. .534

/organism="Phytophthora sojae"

/mol_type="mRNA"

/db_xref="taxon:67593"

/clone_lib="SHB036B09"

/tissue_type="infected host tissue"

/cell_line="P6497"

/dev_stage="48 hour post infection"

/clone_lib="pPHB: Infected hypocotyl soybean host. 48 hrs post infection"

/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; USDA-IFAPs: Expression of Phytophthora sojae genes during infection and propagation."

infection and propagation."

87 t

BASE COUNT 101 a 187 c 159 g 87 t

ORIGIN

Query Match 5.0%; Score 40; DB 14; Length 534;

Best Local Similarity 46.7%; Pred. No. 13;

Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 223 TATGAAGTGGCAACGTGTCGGGATGTACCATGTACGAAAGCTGCTCAACTCAAGC 282

Db 200 TACGGGTGGCGGAGATTACGGTATCCGATGGGTTCCGGCTTCTACAACTGGACC 259

Qy 283 ATTGTGTATGAGCAGCGACATGATCATGTGCACACCCCGGTGGTGCCTCGCTCGG 342

Db 260 TCGATGGACAGAGAGGCGCCCGATCATGCTGACCCCAAGACGGTGGCCAACTTAC 319

Qy 343 GAGAACACTCTTCCCGCTGCTGGGTAGCGCTCACCCCAAGCTGCTAGGACGCC 402

Db 320 CACTACGCGCGCACCATCTCTCGGTGTAACCGTGGTTCGACGTGGACAGATTATC 379

Qy 403 AGCGTCCCAACACAGACATACGACGACCGTTCGATTTGCTGTTGGGCGGCTGTTTC 462

Db 380 AACTTCTGAGCAGAACGCGGTCTCGAGGTATAGTATCGCGGTGACGCGACCCAC 439

Qy 463 TGTTCGGGTATGACGTGGGGACCTCTGCGG 494

Db 440 CGTGGCGCAACAGATCTCGGAGAGTGCGG 471

RESULT 14

BI723733

LOCUS

DEFINITION

Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION

BI723733

VERSION

EST.

KEYWORDS

Chlamydomonas reinhardtii

SOURCE

Chlamydomonas reinhardtii

ORGANISM

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Search completed: December 20, 2003, 06:54:39
Job time : 2050.18 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 1828.77 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATGCCCGGTGCTCTTCTC.....TGATGAAGTGGTCTTAATAG 483

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rnd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	483	100.0	483	6	A48671	A48671 Sequence 9
2	483	100.0	483	6	AR157327	AR157327 Sequence
3	483	100.0	483	6	AX452758	AX452758 Sequence
4	483	100.0	483	6	AX685010	AX685010 Sequence
5	474.2	98.2	480	6	A48673	A48673 Sequence 11
6	474.2	98.2	480	6	AR157328	AR157328 Sequence
7	474.2	98.2	480	6	AX452760	AX452760 Sequence
8	474.2	98.2	480	6	AX685012	AX685012 Sequence
9	438.2	90.7	9379	14	AF207766	AF207766 Hepatitis C
10	435	90.1	1880	14	HPC5TRJ4	D00832 Hepatitis C
11	435	90.1	2540	6	E04260	E04260 cDNA encodi
12	435	90.1	2540	6	E04805	E04805 cDNA to 5'-
13	435	90.1	2540	6	E07391	E07391 cDNA encodi
14	435	90.1	9448	14	HPCJ483	D13558 Hepatitis C
15	433.4	89.7	1539	6	AR027786	AR027786 Sequence
16	433.4	89.7	1863	6	AR027783	AR027783 Sequence
17	433	89.6	9580	14	AF054250	AF054250 Hepatitis
18	431.8	89.4	8780	14	AF054257	AF054257 Hepatitis
19	431.8	89.4	8780	14	AF054258	AF054258 Hepatitis
20	430.2	89.1	8780	14	AF054255	AF054255 Hepatitis
21	430.2	89.1	8781	14	AF054254	AF054254 Hepatitis
22	430.2	89.1	9379	14	AF165052	AF165052 Hepatitis
23	430.2	89.1	9460	14	HPCJ491	D10750 Hepatitis C
24	430.2	89.1	9595	6	AR119832	AR119832 Sequence
25	430.2	89.1	9595	14	AF054247	AF054247 Hepatitis
26	430.2	89.1	9596	14	AF054249	AF054249 Hepatitis
27	430.2	89.1	9599	6	AR119833	AR119833 Sequence
28	428.6	88.7	577	6	E04085	E04085 gDNA encodi
29	428.6	88.7	618	14	HPCJ1F12	D28929 Hepatitis C
30	428.6	88.7	8779	14	AF054251	AF054251 Hepatitis
31	428.6	88.7	8780	14	AF054252	AF054252 Hepatitis
32	428.6	88.7	9377	14	AF207756	AF207756 Hepatitis
33	428.6	88.7	9410	14	HPCJ1R2	D50481 Hepatitis C
34	428.6	88.7	9435	14	AB049093	AB049093 Hepatitis
35	428.6	88.7	9595	14	AF054248	AF054248 Hepatitis
36	427	88.4	1595	14	HPCNSISPF	M74809 Hepatitis C
37	427	88.4	1635	14	HPCNSISPD	M74807 Hepatitis C
38	427	88.4	3296	14	AB008446	AB008446 Hepatitis
39	427	88.4	8780	14	AF054259	AF054259 Hepatitis
40	427	88.4	8781	14	AF054256	AF054256 Hepatitis
41	427	88.4	9361	14	AF483269	AF483269 Hepatitis
42	427	88.4	9369	14	AF165054	AF165054 Hepatitis
43	427	88.4	9379	14	AF165051	AF165051 Hepatitis
44	425.4	88.1	577	6	E04089	E04089 gDNA encodi
45	425.4	88.1	1413	14	HPCRNAP	D00574 Hepatitis C

ALIGNMENTS

RESULT 1
A48671
LOCUS A48671 Sequence 9 from Patent WO9604385. 483 bp DNA linear PAT 07-MAR-1997
DEFINITION A48671
ACCESSION A48671
VERSION A48671.1 GI:2302384
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 483)
Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 9 15-FEB-1996;

Mon Dec 22 13:28:53 2003

INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
source
1..483
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..480
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03131.1"
/db_xref="GI:2302385"
/translation="MPGCGSFPFLALLSLIIPASAEVNRVSGVHVTVNDGNSSSI
VYEAADMIMHTPCVPCVREGNSSRCWALPTPLAARNASVPTTIRRHVDLLVGA
FCSAMVYVDLCCGSVFLVSLFTFSPRRHQTVDQNCISYPYGHVSHRWAMDMNWS"
mat_peptide
1..477
/product="unnamed"
BASE COUNT 85 a 152 c 123 g 123 t
ORIGIN
Query Match 100.0%; Score 483; DB 6; Length 483;
Best Local Similarity 100.0%; Pred. No. 3.1e-113;
Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCGGTTCCTTCTCTATCTCTCTGCGCCCTGCTGCTCTGACCATACCA 60
Db 1 ATGCCGGTTCCTTCTCTATCTCTCTGCGCCCTGCTGCTCTGACCATACCA 60
QY 61 GCTTCCGCTTATGAAGTGCACAAAGTGTCCGGGGTGTACCATGTACGAAACGACTGCTCC 120
Db 61 GCTTCCGCTTATGAAGTGCACAAAGTGTCCGGGGTGTACCATGTACGAAACGACTGCTCC 120
QY 121 AACTCAAGCATAGTGTATGAGGACGCGGACATCATGCACACCCCGGGTGCCTGCC 180
Db 121 AACTCAAGCATAGTGTATGAGGACGCGGACATCATGCACACCCCGGGTGCCTGCC 180
QY 181 TGCGTTCGGGAGGGCAACTCTCTCCGTTGCTGGGTGGCGTCACTCCACGCTCGGGCC 240
Db 181 TGCGTTCGGGAGGGCAACTCTCTCCGTTGCTGGGTGGCGTCACTCCACGCTCGGGCC 240
QY 241 AGGAACGCCAGCGTCCCAACAGCAATACGACGCGCACCGTCCGATTGCTCGTGGGGCT 300
Db 241 AGGAACGCCAGCGTCCCAACAGCAATACGACGCGCACCGTCCGATTGCTCGTGGGGCT 300
QY 301 GCTGCTTTCTGTCGGCTATGTACGTGGGGATCTCTGCGGATCTGTTCTTGTGTTTC 360
Db 301 GCTGCTTTCTGTCGGCTATGTACGTGGGGATCTCTGCGGATCTGTTCTTGTGTTTC 360
QY 361 CAGCTGTTTCACTTCTCACTCCGCGGATCAACAGTACAGGACTGCACTGCTCAATC 420
Db 361 CAGCTGTTTCACTTCTCACTCCGCGGATCAACAGTACAGGACTGCACTGCTCAATC 420
QY 421 TATCCCGGCATGTATCAGGTCAACGATGCTTGGGATATGATGATGAATGCTGCTAA 480
Db 421 TATCCCGGCATGTATCAGGTCAACGATGCTTGGGATATGATGATGAATGCTGCTAA 480
QY 481 TAG 483
Db 481 TAG 483
RESULT 3
AX452758 AX452758 483 bp DNA linear PAT 06-JUL-2002
LOCUS Sequence 9 from Patent EP1211315.
DEFINITION
ACCESSION AX452758
VERSION AX452758.1 GI:21712443
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1
AUTHORS Maertens G., Bosman, F., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 9 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
Location/Qualifiers
1..483
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..480
CDS
RESULT 2
ARI57327
LOCUS Sequence 9 from patent US 6245503.
DEFINITION
ACCESSION ARI57327
VERSION ARI57327.1 GI:16218260
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 483)

[illegible]

FEATURES		Location/Qualifiers	
source	1..1880	/organism="Hepatitis C virus"	
		/mol_type="genomic RNA"	
		/isolate="HC-J4"	
		/db_xref="taxon:11103"	
		/note="82 bp upstream of StyI site"	
CDS	342..>1880		
		/note="putative"	
		/codon_start=1	
		/product="structural protein"	
		/protein_id="BAA0706.1"	
		/db_xref="GI:221514"	
		/translation="MTNPKPQKTRNTRRQDVKFPGGQIVGGVILLPRGPRRL GVNRTKTSRSPQWRQPIPKARPEGRAPQGPMPLYNEGLWAGWLLSPRG SRPSWGTPPRRSRLNKGVIDLTGCFADLMGYIPLVGAFLGGAARALAHGVRLVED GVNATGNLPGGSGFSFLALLSLCLTIPASAYEVNRVSGIYHVNDSCNSIIVYEAD MIMHTPGVPCVEDNSSRCWALTPTLAARNASVPTTIRRHVDLLVGAAPCSAMY VGDLCGSFLVLSQLFTFSPRRHETVDCNCSLYPHLSGHRMAWDMNWSPTTALVV SQLRLTPQAVDMVAGAHGVLGLAYYSVMGNWAKVLIVALLFAGVDGETYSGGAA SHTTSLASLSPGASORIQLVNTGNSWHINRTALNCNDSLHTGFLAALFYTHRFNS GCPERMASCRPIDWFAQGWGPIITYTBPDSPDQRPYCHVYAPRCGIVPASOVCGPYVC FTPSP"	
BASE COUNT	333 a 593 c 555 g 399 t		
ORIGIN			
Query Match 90.1%; Score 435; DB 14; Length 1880;			
Best Local Similarity 94.7%; Pred. No. 6.5e-101;			
Matches 450; Conservative 0; Mismatches 25; Indels 0; Gaps 0;			
QY	2	TGCCCCGGTGTCTTTCTCTATCTTCTCTTGGCCCTGTCTGTCTGTACCATACCAG 61	
DB	847	TGCCCCGGTGTCTTTCTCTATCTTCTCTTGGCCCTGTCTGTCTGTACCATACCAG 906	
QY	62	CTTCCCGCTATGAAGTGGCGCAAGTGTCCGGGTGTACCATGTACGAAAGCTGTCTCA 121	
DB	907	CTTCCCGCTATGAAGTGGCGCAAGTGTCCGGGTGTACCATGTACGAAAGCTGTCTCA 966	
QY	122	ACTCAAGCATAGTGTATGAGGCGAGCGACATGATCATGTACACCCCGGGTGGTGCCT 181	
DB	967	ACTCAAGCATAGTGTATGAGGCGAGCGACATGATCATGTACACCCCGGGTGGTGCCT 1026	
QY	182	GCCTTCGGGAGGCGCAACTCTCCCGTGTGGGTGGCGCTCACTCCACCGCTCGCGCCA 241	
DB	1027	GCCTTCGGGAGGCGCAACAGCTCCCGTGTGGGTGGCGCTCACTCCACCGCTCGCGCCA 1086	
QY	242	GGAAAGCCAGCGTCCCAACACAAATACGACGCGACGTCGATTTGCTGTTGGGGCTG 301	
DB	1087	GGAAAGCCAGCGTCCCAACACAAATACGACGCGACGTCGATTTGCTGTTGGGGCTG 1146	
QY	302	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTTGTTCCTCC 361	
DB	1147	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTTGTTCCTCC 1206	
QY	362	AGCTGTTTACCTTCTCACTCGCGGCGCATCAAAAGTACAGGACTGCAATCTCAATCT 421	
DB	1207	AGCTGTTTACCTTCTCGCTCGCGGCGCATGAGACAGTGCAGGACTGCAATCTCAATCT 1266	
QY	422	ATCCCGGCGCATGTATCAGGTACCGCATGGCTTTGGGATATGATGATGAATGGTCTC 476	
DB	1267	ATCCCGGCGCATTTATCAGGTACCGCATGGCTTTGGGATATGATGATGAATGGTCTC 1321	
RESULT 11			
E04260			
LOCUS	E04260	2540 bp	RNA linear PAT 29-SEP-1997
DEFINITION	CDNA encoding a part of type non-A non-B hepatitis virus.		
ACCESSION	E04260		
VERSION	E04260.1	GI:2172463	
KEYWORDS	JP 1993023200-A/2.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE			
1 (bases 1 to 2540)			
Okamoto,H. and Nakamura,T.			
HIGHLY SENSITIVE DETECTION METHOD OF NON-A NON-B TYPE HEPATITIS			
VIRUS USING OLIGONUCLEOTIDE PRIMER AND OLIGONUCLEOTIDE PRIMER			
Patent: JP 1993023200-A 2 02-FEB-1993;			
NAKAMURA TETSUO			
PN JP 1993023200-A/2			
PD 02-FEB-1993			
PF 26-APR-1991 JP 1991191376			
PR 12-JUN-1990 JP 90P 153402			
PI OKAMOTO HIROAKI, NAKAMURA TETSUO			
PC C12Q1/68,C12N15/51,C12Q1/70;			
CC strandedness: Double;			
CC topology: Linear;			
CC hypothetical: No;			
CC anti-sense: No;			
CC *source: strain=HC-J4;			
FH Key Location/Qualifiers			
FH misc_feature 1..2540			
FT /note='a part of type non-A non-B hepatitis			
FT virus'.			
FT			
FEATURES			
source			
		1..2540	
		/organism="unidentified"	
		/mol_type="genomic RNA"	
		/db_xref="taxon:32644"	
BASE COUNT	470 a 776 c 741 g 553 t		
ORIGIN			
Query Match 90.1%; Score 435; DB 6; Length 2540;			
Best Local Similarity 94.7%; Pred. No. 6.6e-101;			
Matches 450; Conservative 0; Mismatches 25; Indels 0; Gaps 0;			
QY	2	TGCCCCGGTGTCTTTCTCTATCTTCTCTTGGCCCTGTCTGTCTGTACCATACCAG 61	
DB	847	TGCCCCGGTGTCTTTCTCTATCTTCTCTTGGCCCTGTCTGTCTGTACCATACCAG 906	
QY	62	CTTCCCGCTATGAAGTGGCGCAAGTGTCCGGGTGTACCATGTACGAAAGCTGTCTCA 121	
DB	907	CTTCCCGCTATGAAGTGGCGCAAGTGTCCGGGTGTACCATGTACGAAAGCTGTCTCA 966	
QY	122	ACTCAAGCATAGTGTATGAGGCGAGCGACATGATCATGTACACCCCGGGTGGTGCCT 181	
DB	967	ACTCAAGCATAGTGTATGAGGCGAGCGACATGATCATGTACACCCCGGGTGGTGCCT 1026	
QY	182	GCCTTCGGGAGGCGCAACTCTCCCGTGTGGGTGGCGCTCACTCCACCGCTCGCGCCA 241	
DB	1027	GCCTTCGGGAGGCGCAACAGCTCCCGTGTGGGTGGCGCTCACTCCACCGCTCGCGCCA 1086	
QY	242	GGAAAGCCAGCGTCCCAACACAAATACGACGCGACGTCGATTTGCTGTTGGGGCTG 301	
DB	1087	GGAAAGCCAGCGTCCCAACACAAATACGACGCGACGTCGATTTGCTGTTGGGGCTG 1146	
QY	302	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTTGTTCCTCC 361	
DB	1147	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTTGTTCCTCC 1206	
QY	362	AGCTGTTTACCTTCTCACTCGCGGCGCATCAAAAGTACAGGACTGCAATCTCAATCT 421	
DB	1207	AGCTGTTTACCTTCTCGCTCGCGGCGCATGAGACAGTGCAGGACTGCAATCTCAATCT 1266	
QY	422	ATCCCGGCGCATGTATCAGGTACCGCATGGCTTTGGGATATGATGATGAATGGTCTC 476	
DB	1267	ATCCCGGCGCATTTATCAGGTACCGCATGGCTTTGGGATATGATGATGAATGGTCTC 1321	
RESULT 12			
E04805			
LOCUS	E04805	2540 bp	RNA linear PAT 29-SEP-1997
DEFINITION	CDNA to 5'-terminal region of grNA of Hepatitis nonA nonB virus.		
ACCESSION	E04805		
VERSION	E04805.1	GI:2173001	

KEYWORDS	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
REFERENCE	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
AUTHORS	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
TITLE	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
JOURNAL	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
COMMENT	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
REFERENCE	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
AUTHORS	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
TITLE	JP 1993091884-A/2.	unidentified	unidentified	unclassified.	1 (bases 1 to 2540)	Okamoto, H. and Nakamura, T.	OKAMOTO HIROAKI, NAKAMURA TETSUO	PC C12N15/51, C07K13/00, C12N15/06, C12P21/08, G01N33/53, G01N33/576, G01N33/577;	CC topology: Linear;	CC hypothetical: No;	CC anti-sense: No;	Location/Qualifiers	Key	5'UTR	342..2540	/note="protein coded by Hepatitis nonA nonB virus gene".	Location/Qualifiers	1..2540	/organism="unidentified"	/mol_type="genomic RNA"	/db_xref="taxon:32644"	470 a	776 c	741 g	553 t	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	25;	Indels	0;	Gaps	0;
JOURNAL	JP 1993091884-A/2.	unidentified																																				

Db	806	CTGCTTTCTGCTCCGCTATGTACGTGGGGGATCTCGGGATCTGTTTCTCGTGNCTCC	865
Qy	362	AGTGTTCACTTCTCACCTCCGGGATCAACAGTACAGGACTGCAACTGCTCAATCT	421
Db	866	AGCTGTTCACTTCTCGGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCTCAATCT	925
Qy	422	ATCCGGCCATGATACAGTCAACCGCATGGCTGGGATATGATGAACTGTC	476
Db	926	ATCCGGCCATTTATCAGGTACCCGATGGCTTGGGATATGATGAACTGGTC	980

Search completed: December 20, 2003, 02:01:52
Job time : 1830.77 secs

QY	122	ACTCAAGCATGATGTATGAGCGAGCAGCATGATCATGCACACCCTCCGGGTGCGTGCCT	181
DB	967	ACTCAAGCATTTGTTATGAGCAGCGGACATGATCATGTCATCTCCGGGTGCGTGCCT	1026
QY	182	GCGTTGGGAGGCCAACTCCTCCCGTTGCTGGGTGGGCTCACTCCACGCTCGCGGCA	241
DB	1027	GCGTTGGGAGGACAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGGCA	1086
QY	242	GGAAAGCAGCGTCCCACAAACGCAATACGACGCACGTCGAATTGCTCGTTGGGGTG	301
DB	1087	GGAAAGCAGCGTCCCACACACGCAATACGACGCACGTCGAATTGCTCGTTGGGGTG	1146
QY	302	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTTGGCGATCTGTTTCTGTTTCCC	361
DB	1147	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTTGGCGATCTGTTTCTGTTTCCC	1206
QY	362	AGCTGTTCACTTCTCACCTCGCGGCATCAAACAGTACAGGACTGCAACTGCTCAATCT	421
DB	1207	AGCTGTTCACTTCTCGCGCTCGCGGCATGACAGTGTGAGGACTGCAACTGCTCAATCT	1266
QY	422	ATCCCGGCATGTATCAGGTCAACGCATGCTTGGGATATGATGAACTGTC	476
DB	1267	ATCCCGGCATTTATCAGGTCAACGCATGCTTGGGATATGATGAACTGTC	1321
RESULT 15	AR027786	1539 bp DNA linear PAT 29-SEP-1999	
LOCUS	Sequence 17 from patent US 5856458.		
DEFINITION	AR027786		
ACCESSION	AR027786.1 GI:5938606		
VERSION			
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1539)		
AUTHORS	Okamoto,H. and Nakamura,T.		
TITLE	Oligonucleotide primers, and their application for high-fidelity detection of non-A, non-B hepatitis virus		
JOURNAL	Patent: US 5856459-A 17 05-JAN-1999;		
FEATURES	Location/Qualifiers		
source	1..1539 /organism="unknown"		
BASE COUNT	271 a 490 c 448 g 330 t		
ORIGIN			
Query Match	89.7%; Score 433.4; DB 6; Length 1539;		
Best Local Similarity	94.5%; Pred. No. 1.7e-100;		
Matches 449; Conservative	0; Mismatches 26; Indels 0; Gaps 0		
QY	2	TGCCCGGTGCTTTCTCTATCTCTCTCTGCGCCCTGCTGCTGACCATACGAG 61	
DB	506	TGCCCGGTGCTTTCTCTATCTCTCTCTGCGCCCTGCTGCTGACCATACGAG 565	
QY	62	CTTCGCTTTAAGTAGTGGCAACGTGTCCGGGTGTACCATGTACGAAAGCTGTCCA 121	
DB	566	CTTCGCTTTAAGTAGTGGCAACGTGTCCGGGTGTACCATGTACGAAAGCTGTCCA 625	
QY	122	ACTCAAGCATGTGTATGAGCAGCGGACATGATCATGACACCCCGGGTGGTGGCCT 181	
DB	626	ACTCAAGCATTTGTATGAGCAGCGGACATGATCATGACATCTCCGGGTGGTGGCCT 685	
QY	182	GCGTTCCGAGGGCAACTCTCCCGTTGCTGGGTGGGCTCATCTCCACGCTCGGGCCA 241	
DB	686	GCGTTCCGAGGGCAACAGCTCCCCGTGCTGGGTAGCGCTCACTCCACGCTCGGGCCA 749	
QY	242	GGAAAGCAGCGTCCCACAAACGCAATACGACGCCAGCTGATTTGCTTGGGGTG 301	
DB	746	GGAAAGCAGCGGTCCCACATACGCAATACGACGCCAGCTGATTTGCTTGGGGTG 805	
QY	302	CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATGATGAACTGTTCCC 361	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 133.529 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-9
Perfect score: 483
Sequence: 1 ATGCCGGTGTCTTCTC.....TGATGAACGTGCTCCTAATAG 483

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483	100.0	483	17	AAT12707
2	483	100.0	483	24	AA148916
3	474.2	98.2	480	17	AAT12708
4	474.2	98.2	480	24	AA148917
5	435	90.1	1880	13	AAQ24467
6	435	90.1	2187	19	ABAQ3491
7	435	90.1	2540	14	AAQ43889
8	435	90.1	2540	15	AAQ63753

ALIGNMENTS

RESULT 1

AAT12707

ID AAT12707 standard; DNA; 483 BP.

XX AC AAT12707;

XX AC AAT12707;

XX DT 23-SEP-1996 (first entry)

XX DE HCV E1 construct HCC112A.

XX DE HCV E1 construct HCC112A.

XX KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

XX KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX KW ss.

XX OS Hepatitis C virus.

XX OS Hepatitis C virus.

XX DN WQ9604385-A2.

XX PD 15-FEB-1996.

XX PP 31-JUL-1995; 95WO-EP03031.

XX PR 29-JUL-1994; 94EP-0870132.

XX XX (INNO-) INNOGENETICS NV.

XX PA Bosman F, Buyse M, De Martynoff G, Maertens G;

XX PI WPI; 1996-129401/13.

XX DR Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

XX PT

Fragment of NANB h
Hepatitis C virus
Infectious hepatitis
Nucleotide sequenc
Infectious hepatitis
Infectious hepatitis
HCV envelope regio
Cuticle protein 1
Fragment of hepati
HCV J1 E domain co
HCV envelope regio
Hepatitis C virus
Consensus sequence
Consensus sequence
Hepatitis C virus
Hepatitis C virus
Consensus sequence
Fragment #5 isolat
Plasmid pDK2 DNA
Fragment #6 isolat
Non-A, non-B hepat
5'UTR/CORE/ENV/NS1
Non-A, non-B hepat
5'UTR/CORE/ENV/NS1
NANB hepatitis vir
HCV E1 construct H
Hepatitis C virus
Fragment #1 isolat
HCV envelope regio
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
HCV E1 construct H
Hepatitis C virus
Hepatitis C virus
NANB hepatitis vir
Hepatitis C virus

proteins - in presence of di:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
SQ Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;
Query Match 100.0%; Score 483; DB 17; Length 483;
PT Best Local Similarity 100.0%; Pred. No. 2.6e-131; Indels 0; Gaps 0;
PT Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ATGCCCGGTGCTCTTTCTATCTCTCTTCTGCGCTGCTGCTCTGTCGACCATACCA 60
Db 1 ATGCCCGGTGCTCTTTCTATCTCTCTTCTGCGCTGCTGCTCTGTCGACCATACCA 60
QY 61 GCTTCCGCTTATGAAGTGGCGAACGTGTCGGGGTGACCATGTCCAGACGCTGCTCC 120
Db 61 GCTTCCGCTTATGAAGTGGCGAACGTGTCGGGGTGACCATGTCCAGACGCTGCTCC 120
QY 121 AACTCAAGCATAGTGTATGAGCGCGGACATGATCATGACACCCCGGGTGGCTGCC 180
Db 121 AACTCAAGCATAGTGTATGAGCGCGGACATGATCATGACACCCCGGGTGGCTGCC 180
QY 181 TGGCTTCCGAGGCGCAACTCTCCCGTGGTGGCGCTCACTCCACGCTCGCGGCC 240
Db 181 TGGCTTCCGAGGCGCAACTCTCCCGTGGTGGCGCTCACTCCACGCTCGCGGCC 240
QY 241 AGGAACGCCAGCGTCCCGCATCAACAGTACAGACGCTGCAACTGCTCAATC 420
Db 241 AGGAACGCCAGCGTCCCGCATCAACAGTACAGACGCTGCAACTGCTCAATC 420
QY 301 GCTGCTTTCTGTTCCGCTATGATACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
Db 301 GCTGCTTTCTGTTCCGCTATGATACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
QY 361 CAGGTGTTCACTTCTCACTCCGCGGCATCAACAGTACAGACGCTGCAACTGCTCAATC 420
Db 361 CAGGTGTTCACTTCTCACTCCGCGGCATCAACAGTACAGACGCTGCAACTGCTCAATC 420
QY 421 TATCCCGGCGCATGATCAGGTTCACCGCATGGCTTGGGATATGATGATGAACTGGTCTTAA 480
Db 421 TATCCCGGCGCATGATCAGGTTCACCGCATGGCTTGGGATATGATGATGAACTGGTCTTAA 480
QY 481 TAG 483
Db 481 TAG 483
XX
AC AAL48916 standard; DNA; 483 BP.
XX AAL48916;
AC AAL48916;

24-OCT-2002 (first entry)
Hepatitis C virus clone HCC112A E1 protein coding sequence.
Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
virucide; immunostimulant; vaccine; ds.
Hepatitis C virus.
WO200255548-A2.
18-JUL-2002.
11-JAN-2002; 2002WO-EP00219.
11-JAN-2001; 2001US-260699P.
30-AUG-2001; 2001US-315768P.
(INNO-) INNOGENETICS NV.
Maertens G, Bosman F, Buyse M;
WPI; 2002-599657/64.
P-PSDB; MAO18663.
New therapeutic vaccine compositions comprising at least one purified
recombinant hepatitis C virus (HCV) single or specific oligomeric
recombinant envelope protein E1 or E2, useful for immunizing humans
from HCV infection -
Example 2; Page 165-166; 243pp; English.
The present invention relates to new therapeutic vaccine compositions for
inducing hepatitis C virus (HCV)-specific antibodies, comprising a
composition containing at least one purified recombinant HCV single or
specific oligomeric recombinant envelope proteins selected from an E1 and
an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
useful for inducing HCV-specific antibodies or for immunising humans
against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
vaccines or therapeutics, in HCV screening and confirmatory antibody
tests, for raising antibodies, in the preparation of medication, and for
in vitro monitoring of HCV disease or prognosis of the response to
treatment of patients suffering from HCV infection. The present sequence
is a coding sequence described in the exemplification of the invention.
Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;
Query Match 100.0%; Score 483; DB 24; Length 483;
PT Best Local Similarity 100.0%; Pred. No. 2.6e-131; Indels 0; Gaps 0;
PT Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCCCGGTGCTCTTTCTATCTCTCTTCTGCGCTGCTGCTCTGTCGACCATACCA 60
Db 1 ATGCCCGGTGCTCTTTCTATCTCTCTTCTGCGCTGCTGCTCTGTCGACCATACCA 60
QY 61 GCTTCCGCTTATGAAGTGGCGAACGTGTCGGGGTGACCATGTCCAGACGCTGCTCC 120
Db 61 GCTTCCGCTTATGAAGTGGCGAACGTGTCGGGGTGACCATGTCCAGACGCTGCTCC 120
QY 121 AACTCAAGCATAGTGTATGAGCGCGGACATGATCATGACACCCCGGGTGGCTGCC 180
Db 121 AACTCAAGCATAGTGTATGAGCGCGGACATGATCATGACACCCCGGGTGGCTGCC 180
QY 181 TGGCTTCCGAGGCGCAACTCTCCCGTGGTGGCGCTCACTCCACGCTCGCGGCC 240
Db 181 TGGCTTCCGAGGCGCAACTCTCCCGTGGTGGCGCTCACTCCACGCTCGCGGCC 240
QY 241 AGGAACGCCAGCGTCCCGCATCAACAGTACAGACGCTGCAACTGCTCAATC 300
Db 241 AGGAACGCCAGCGTCCCGCATCAACAGTACAGACGCTGCAACTGCTCAATC 300
QY 301 GCTGCTTTCTGTTCCGCTATGATACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
Db 301 GCTGCTTTCTGTTCCGCTATGATACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCC 360
XX
SQ

122 ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGCACACCCCGGGTGCCT 181
 967 ACTCAAGCATGTGTATGAGGCGAGCGGACATGATCATGCTACTCCGGGTGGTGCCT 1026
 182 GCCTTCGGGAGGCAACTCTCCCTGTTGCTGGGTGGGTGCTCACTCCAGCTCGGGCCA 241
 1027 GCCTTCGGGAGGCAACTCTCCCTGTTGCTGGGTGGGTGCTCACTCCAGCTCGGGCCA 1086
 242 GGAAGCGCAGCTGCTCCCAACAGCAATAGCAGCGCACTGCAATTTGCTGGGGCTG 301
 1087 GGAATGCGAGCTGCTCCCAACAGCAATAGCAGCGCACTGCAATTTGCTGGGGCTG 1146
 362 AGCTGTTTCACTTCTACCTGCGCGCATCAACAGTACAGGACTGCACTGCTCAATCT 421
 1207 AGCTGTTTCACTTCTACCTGCGCGCATCAACAGTACAGGACTGCACTGCTCAATCT 1266
 422 ATCCGGGCGCATGATCAGTCAACCGCATGGCTTGGGATATGATGAACTGGTC 476
 1267 ATCCGGGCGCATGATCAGTCAACCGCATGGCTTGGGATATGATGAACTGGTC 1321

RESULT 8
 AAQ63753
 ID AAQ63753 standard; cDNA to mRNA; 2540 BP.
 AC AAQ63753;
 AC AAQ63753;
 25-MAR-2003 (updated)
 17-MAR-1992 (first entry)
 Fragment of NANB hepatitis virus strain HC-J4.
 Post-transfusion; non-A, non-B hepatitis; HCV; detection primer; PCR;
 polymerase chain reaction; ss.
 Hepatitis C virus.
 Key Location/Qualifiers
 5'UTR 1..324
 CDS /*tag= a
 325..1863
 /*tag= b
 EP461863-A.
 PN
 XX
 XX
 PF 18-DEC-1991.
 XX
 PF 11-JUN-1991; 91EP-0305270.
 XX
 PR 12-JUN-1990; 90JP-0153402.
 XX
 PA (IMMO) IMMUNO JAPAN INC.
 XX
 PI Okamoto H, Ogikubo Y, Nakamura T;
 XX
 WPI; 1991-370834/51.
 XX
 XX
 PT Oligo-nucleotide primers - derived from and used to detect and
 diagnose non-A, non-B hepatitis virus
 XX
 PS Claim 1; Page 7; 23pp; English.
 XX

122 ACTCAAGCATAGTGTATGAGGCGAGCGGACATGATCATGCACACCCCGGGTGCCT 181
 967 ACTCAAGCATGTGTATGAGGCGAGCGGACATGATCATGCTACTCCGGGTGGTGCCT 1026
 182 GCCTTCGGGAGGCAACTCTCCCTGTTGCTGGGTGGGTGCTCACTCCAGCTCGGGCCA 241
 1027 GCCTTCGGGAGGCAACTCTCCCTGTTGCTGGGTGGGTGCTCACTCCAGCTCGGGCCA 1086
 242 GGAAGCGCAGCTGCTCCCAACAGCAATAGCAGCGCACTGCAATTTGCTGGGGCTG 301
 1087 GGAATGCGAGCTGCTCCCAACAGCAATAGCAGCGCACTGCAATTTGCTGGGGCTG 1146
 362 AGCTGTTTCACTTCTACCTGCGCGCATCAACAGTACAGGACTGCACTGCTCAATCT 421
 1207 AGCTGTTTCACTTCTACCTGCGCGCATCAACAGTACAGGACTGCACTGCTCAATCT 1266
 422 ATCCGGGCGCATGATCAGTCAACCGCATGGCTTGGGATATGATGAACTGGTC 476
 1267 ATCCGGGCGCATGATCAGTCAACCGCATGGCTTGGGATATGATGAACTGGTC 1321

RESULT 8
 AAQ63753
 ID AAQ63753 standard; cDNA to mRNA; 2540 BP.
 AC AAQ63753;
 AC AAQ63753;
 30-JAN-1995 (first entry)
 NANBHV genomic fragment #2.
 Polymerase chain reaction; PCR; primer; amplify; detection; NANBHV;
 non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
 Synthetic.
 JP06125777-A.
 10-MAY-1994.
 20-JUN-1991; 91JP-0247120.
 20-JUN-1991; 91JP-0247120.
 (NAKA/) NAKAMURA T.
 WPI; 1994-187937/23.
 Oligonucleotide primer pairs specific for non-A, non-B hepatitis
 virus - used for high sensitivity detection of non-A non-B (NANB)
 hepatitis virus
 Disclosure; Page 24-25; 25pp; Japanese.
 The sequences given in AAQ63752-53 represent fragments of the non-A,
 non-B hepatitis virus (NANBH) genome. These fragments were amplified
 using the primers given in AAQ63732-51. These primers were used in the
 detection of NANBH. The primers are based on the 5'-terminal region and
 the core protein coding region. The method allows highly sensitive
 detection of NANBH.
 Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;
 Query Match 90.1%; Score 435; DB 15; Length 2540;
 Best Local Similarity 94.7%; Pred. No. 5.2e-117;
 Matches 450; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
 2 TGCCCGGTGCTCTTTCTCTATCTTCTCTTGGCCCTGCTCTCTGCTGACCATACCAG 61

Query Match 90.1%; Score 435; DB 15; Length 2540;
 Best Local Similarity 94.7%; Pred. No. 5.2e-117;
 Matches 450; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
 2 TGCCCGGTGCTCTTTCTCTATCTTCTCTTGGCCCTGCTCTCTGCTGACCATACCAG 61

CC NANB hepatitis virus strain HC-J4 was isolated from a plasma sample
CC of a chimpanzee challenged with NANB hepatitis for infectivity but
CC which tested negative for HCV antibody by Ortho HCV Ab ELISA test.
CC RNA was isolated from the sample and reverse transcribed into cDNA.
CC The 513 amino acids encoded by the CDS were determined but are not
CC given in the specification (and hence are not included in A-Geneseq).
CC A study of the deduced sequence suggested that the CDS encodes NANBH
CC virus core proteins. Primers for detecting NANB hepatitis virus were
CC designed based on the HC-J4 sequence.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 1863 BP; 333 A; 586 C; 547 G; 397 T; 0 other;

Query Match 89.7%; Score 433.4; DB 12; Length 1863;
Best Local Similarity 94.5%; Pred. No. 1.4e-116;
Matches 449; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCCCTGCTGCTGCTGACCATACCG 61
Db |||||
QY 830 TGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCTTGTGCTGTTGACCATCCAG 889
Db |||||
QY 62 CTTCCGCTTATGAAGTGCACCAAGTGTCCGGGTGTACCATGTACGAAAGACTGTCTCCA 121
Db |||||
QY 890 CTTCCGCTTATGAAGTGCACCAAGTGTCCGGGTGTACCATGTACGAAAGACTGTCTCCA 949
Db |||||
QY 122 ACTCAAGCATAGTGTATGAGGACGCGGACATGATCATGCATACATCCCGGGTGGTGCCT 181
Db |||||
QY 950 ACTCAAGCATAGTGTATGAGGACGCGGACATGATCATGCATACATCCCGGGTGGTGCCT 1009
Db |||||
QY 182 GCGTTCCGGAGGCAACTCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGCCA 241
Db |||||
QY 1010 GCGTTCCGGAGGCAACAGCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGCCA 1069
Db |||||
QY 242 GGAACGCGACGTCCTCCACAAACACATACGACGCACTCGATTCTGTTGGGCTG 301
Db |||||
QY 1070 GGAATGCCAGGTCCTCCACATACGACATACGACGCACTCGATTCTGTTGGGCGG 1129
Db |||||
QY 302 CTGCTTTCTGTCGGTATGTAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361
Db |||||
QY 1130 CTGCTTTCTGTCGGTATGTAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1189
Db |||||
QY 362 AGCTGTTCACTTCTCACTCGCGGATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db |||||
QY 1190 AGCTGTTCACTTCTCGCTCGCGGATGAGACAGTGCAGACTGCAACTGCTCAATCT 1249
Db |||||
QY 422 ATCCCGGCAATATCAGGTACCGCATGCTGGGATATGATGATGAATGGTTC 476
Db |||||
QY 1250 ATCCCGGCAATATCAGGTACCGCATGCTGGGATATGATGATGAATGGTTC 1304
Db |||||

RESULT 10
AAQ29628
ID AAQ29628 standard; DNA; 2540 BP.
XX
AC AAQ29628;
XX
DT 25-MAR-2003 (updated)
DT 16-MAR-1993 (first entry)
XX
DE Hepatitis C virus HC-J4 5' region.
XX
KW Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
KW PCR; primer; polymerase chain reaction; ss.
XX
OS Hepatitis C virus.
XX
PN EP510952-A1.
XX
PD 28-OCT-1992.
XX
PF 23-APR-1992; 92EP-0303625.
XX

PR 26-APR-1991; 91JP-0191376.
XX
PA (IMMO) IMMUNO JAPAN INC.
XX
PI Nakamura T, Okamoto H;
XX
XX WPI; 1992-359137/44.
XX
PT Detection of non-A, non-B hepatitis virus - using new
PT oligo-nucleotide primers with nucleotide sequences corresp. to
PT part. of the viral RNA
XX
PS Disclosure; Page 18; 54pp; English.
XX
CC This sequence represents the 5' region of hepatitis C virus RNA. The
CC original sample was obtained from human and chimpanzee plasma. RNA
CC was isolated from several samples and homology compared, and the
CC respective sequence of about 1900 - 2500 nucleotides of the 5'
CC terminus and 1100 nucleotides of the 3' terminus determined. The 5'
CC region (given) contains a non-coding region of at least 340
CC nucleotides and a region coding for the structural protein followed
CC by a region coding for the non-structural protein (none actually
CC detailed on the sequence given in the specification). When compared
CC with the sequence of HCV disclosed in EP-388232 this sequence showed
CC homology of 80.5%.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2540 BP; 472 A; 775 C; 741 G; 552 T; 0 other;

Query Match 89.4%; Score 431.8; DB 13; Length 2540;
Best Local Similarity 94.3%; Pred. No. 4.5e-116;
Matches 448; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCCCTGCTGCTGCTGACCATACCG 61
Db |||||
QY 847 TGCCCGGTTGCTCTTTCTCTATCTTCTCATGCTTGTGCTGTTGACCATCCAG 906
Db |||||
QY 62 CTTCCGCTTATGAAGTGCACCAAGTGTCCGGGTGTACCATGTACGAAAGACTGTCTCCA 121
Db |||||
QY 907 CTTCCGCTTATGAAGTGCACCAAGTGTCCGGGTGTACCATGTACGAAAGACTGTCTCCA 966
Db |||||
QY 122 ACTCAAGCATAGTGTATGAGGACGCGGACATGATCATGCATACATCCCGGGTGGTGCCT 181
Db |||||
QY 967 ACTCAAGCATAGTGTATGAGGACGCGGACATGATCATGCATACATCCCGGGTGGTGCCT 1026
Db |||||
QY 182 GCGTTCCGGAGGCAACTCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGCCA 241
Db |||||
QY 1027 GCGTTCCGGAGGCAACAGCTCCGTTGCTGGGTAGGCTCACTCCCAAGCTCGCGCCA 1086
Db |||||
QY 242 GGAACGCGACGTCCTCCACAAACACATACGACGCACTCGATTCTGTTGGGCTG 301
Db |||||
QY 1087 GGAATGCCAGGTCCTCCACATACGACATACGACGCACTCGATTCTGTTGGGCGG 1146
Db |||||
QY 302 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361
Db |||||
QY 1147 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1206
Db |||||
QY 362 AGCTGTTCACTTCTCACTCGCGGATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db |||||
QY 1207 AGCTGTTCACTTCTCGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
Db |||||
QY 422 ATCCCGGCAATATCAGGTACCGCATGCTGGGATATGATGATGAATGGTTC 476
Db |||||
QY 1267 ATCCCGGCAATATCAGGTACCGCATGCTGGGATATGATGATGAATGGTTC 1321
Db |||||

RESULT 11
AAQ24843
ID AAQ24843 standard; DNA; 9595 BP.
XX
AC AAQ24843;
XX
DT 21-JUN-1999 (first entry)

XX Infectious hepatitis C virus genotype 1b strain HC-J4 genome.
 XX HCV; infectious clone; infection; diagnosis; therapy; vaccine;
 XX screening; assay; antiviral; virucide; ss.
 XX Hepatitis C virus.
 XX Key Location/Qualifiers
 XX CDS 342..9374
 XX /*tag= a
 XX WO9904008-A2.
 XX 28-JAN-1999.
 XX 16-JUL-1998; 98WO-US14688.
 XX 27-JAN-1998; 98US-0014416.
 XX 18-JUL-1997; 97US-0053062.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Bukh J, Emerson SU, Purcell RH, Yanagi M;
 XX WPI: 1999-132252/11.
 XX P-PSDB; AAW98022.
 XX New isolated hepatitis C virus nucleic acids - used to develop
 XX products for the diagnosis, prevention and treatment of HCV
 XX infections and for developing screening assays
 XX Claim 3; Fig 14A-F; 126pp; English.
 XX The present sequence comprises the nucleic acid sequence of the
 XX genome of infectious hepatitis C virus (HCV) genotype 1b strain
 XX HC-J4 (ATCC 209596) that is capable of expressing this virus when
 XX transfected into cells. HC-J4 was obtained from acute phase plasma
 XX of a chimpanzee experimentally infected with serum containing
 XX HC-J4/91. The claimed infectious nucleic acid sequence can be used
 XX to produce chimeric genomes (see AAX24833) consisting of the open
 XX reading frames of infectious nucleic acid sequences of other
 XX genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
 XX 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
 XX the introduction of mutations or deletions into infectious nucleic
 XX acid sequences in order to produce an attenuated HCV virus suitable
 XX for vaccine development. Infectious nucleic acid sequences can
 XX also be used to produce attenuated virus via passage in vitro or in
 XX vivo of the viruses produced by transfection of a host cell with
 XX the infectious nucleic acid sequence. Vaccines comprising one or
 XX more polypeptides made from the infectious nucleic acid sequence are
 XX used to immunise mammals, especially humans, against hepatitis C.
 XX The nucleic acid sequences can also be used to induce protective
 XX immunity against the virus. The nucleic acid sequences or their
 XX encoded proteases (e.g. NS3 protease) can additionally be used to
 XX develop screening assays to identify antiviral agents for HCV.
 XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;
 XX Query Match 89.1%; Score 430.2; DB 20; Length 9595;
 XX Best Local Similarity 94.1%; Pred. No. 2.1e-115;
 XX Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 XX 2 TGCCGGTGTCTTTCTATCTCTCTTGGCCCTGCTCTCTGACCATACCAAG 61
 XX 847 TGCCGGTGTCTTTCTATCTCTCTTGGCCCTGCTCTCTGACCATACCAAG 906
 XX 62 CTTCCGCTTAAAGTGGCAACGTGTCGGGGGTATACCATGTGACGACCACTGTCTCA 121
 XX 907 CTTCCGCTTAAAGTGGCAACGTGTCGGGGGTATACCATGTGACGACCACTGTCTCA 966
 XX 122 ACTCAAGCATGTATGAGGACGACCATGATCATGCACACCCCGGTGGTGCCT 181

Db 967 ACTCAGCATTGTATGAGGACGAGCGTGTATCATGCATATCCCGGGTGGTGCCT 1026
 Qy 182 GCCTTCGGAGGCGCAACTCTCCCGTTGCTGGGTGGGTCTACTCCACGCTCGCGCCA 241
 Db 1027 GTGTTTCAGAGGGGTAAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCAGCTCGCGCCA 1086
 Qy 242 GGAACGCGCGTCCCGCACACGACATACGACGCGCACGTCGATTTGCTGTTGGGGCTG 301
 Db 1087 GGAATGCCAGCGTCCCGCACACGACATACGACGCGCACGTCGATTTGCTGTTGGGGCTG 1146
 Qy 302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTGTTTCCC 361
 Db 1147 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTGTTTCCC 1206
 Qy 362 AGCTGTTTCACTTCTCACTCCCGCGCATCAAAACAGTACAGGACGTGCAACTGCTCAATCT 421
 Db 1207 AGCTGTTTCACTTCTCGCTTCGCGCATGACAGTACAGGACGTGCAACTGCTCAATCT 1266
 Qy 422 ATCCCGGCGCATGTATCAGGTACCGCATGCGTGGGATATGATGATGAATGCTGTC 476
 Db 1267 ATCCCGGCGCATGTATCAGGTACCGCATGCGTGGGATATGATGATGAATGCTGTC 1321
 RESULT 12
 AAC86939
 ID AAC86939 standard; DNA; 9595 BP.
 XX AC AAC86939;
 XX 02-APR-2001 (first entry)
 DE Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.
 XX Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
 XX HCV; vaccine; viral inhibitor; antiviral; ss.
 XX Hepatitis C virus.
 XX Key Location/Qualifiers
 XX CDS 342..9374
 XX /*tag= a
 XX WO200075352-A2.
 XX 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-US15527.
 XX 04-JUN-1999; 99US-0137817.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nam J, Bukh J, Emerson SU, Purcell RH;
 XX WPI: 2001-071081/08.
 XX P-PSDB; AAB31170.
 XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus
 XX genome in which the (non-)structural region has been replaced by
 XX hepatitis C virus (HCV) genome useful for treating or preventing HCV
 XX signs and symptoms -
 XX Disclosure; Fig 4A-F; 97pp; English.
 XX The specification describes a nucleic acid comprising a chimeric virus
 XX genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
 XX the (non-)structural region has been replaced by the (non-)structural
 XX region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
 XX the chimeric virus and the chimeric virus are useful for identifying
 XX cell lines capable of supporting the replication of these chimeric
 XX viruses, in screening for neutralizing antibodies to HCV of different
 XX genotypes, in the production of HCV-BVDV virions, for the development
 XX of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,

CC in studying the molecular properties of HCV indirectly in vitro, and in
 CC identifying inhibitors of viral enzyme activity which would be useful
 CC as antiviral agents. Formulations or compositions comprising the
 CC chimeric virions may be used to treat or prevent the signs and symptoms
 CC of HCV. The present sequence represents a HCV clone, which is used
 CC to construct chimeric nucleic acids of the invention.

XX
 SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Query Match 89.1%; Score 430.2; DB 22; Length 9595;
 Best Local Similarity 94.1%; Pred. No. 2.1e-115;
 Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGTGACCATATCCAG 61
 DB 847 TGCCCGGTTGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGTGACCATATCCAG 906
 QY 62 CTTCCCGTTATGAAGTCGCAACGTGTCGCGGGGTACCATGTACCAACGACTGCTCCA 121
 DB 907 CTTCCCGTTATGAAGTCGCAACGTGTCGCGGGATATACCATGTACCAACGACTGCTCCA 966
 QY 122 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGCGCCT 181
 DB 967 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGCGCCT 1026
 QY 182 GCGTTCCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTCATCTCCACCGTCGCGGCCA 241
 DB 1027 GTGTTTCAGAGGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACCGTCGCGGCCA 1086
 QY 242 GGAAGCGCAGCGTCCCAACGACATACGACGCGACGTCGATTCGCTGTTGGGCTG 301
 DB 1087 GGAATGCCAGCGTCCCAACGACATACGACGCGACGTCGATTCGCTGTTGGGACGG 1146
 QY 302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTGTTTCCC 361
 DB 1147 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTATTTTCTGCTCCC 1206
 QY 362 AGCTGTTCACTTCTCACTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
 DB 1207 AGCTGTTCACTTCTCGCTCGCGCGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
 QY 422 ATCCCGCCCATGTATCAGGTCACCGCATGCGCTTGGGATATGATGAACTGGTC 476
 DB 1267 ATCCCGCCCATGTATCAGGTCACCGCATGCGCTTGGGATATGATGAACTGGTC 1321

RESULT 13

AAF23492
 ID AAF23492 standard; DNA; 9595 BP.

XX
 AC AAF23492;

XX
 DT 21-MAR-2001 (first entry)

XX
 DE Infectious Hepatitis C virus 1b genotype.

XX
 KW GBV-B; hepatitis C virus; HCV; vaccine; ds.

XX
 OS Hepatitis C virus.

XX
 PN WO200075337-A1.

XX
 PD 14-DEC-2000.

XX
 PF 02-JUN-2000; 2000WO-US15293.

XX
 PR 04-JUN-1999; 99US-0137694.

XX
 PA (USSS) US DEPT HEALTH & HUMAN SERVICES.

XX
 PI Bukh J, Yanagi M, Emerson SU, Purcell RH;

XX
 DR WPI; 2001-091214/10.

XX

PT New infectious nucleic acids of the GB virus-B clone, useful for
 PT indirectly studying the molecular properties of hepatitis C virus (HCV)
 PT and in developing vaccines and therapeutics for HCV .

XX
 PS Disclosure; Fig 7; 96pp; English.

XX
 CC The present invention relates to GB virus-B. The nucleic acid molecules
 CC of the invention are useful for indirectly studying the molecular
 CC properties of hepatitis C virus (HCV) . The infectious nucleic acid
 CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
 CC in the development of vaccines and therapeutics for HCV.

XX
 SQ Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 other;

Query Match 89.1%; Score 430.2; DB 22; Length 9595;
 Best Local Similarity 94.1%; Pred. No. 2.1e-115;
 Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2 TGCCCGGTTGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGTGACCATATCCAG 61
 DB 847 TGCCCGGTTGCTCTTCTCTATCTTCTTGGCCCTGCTGCTGTGACCATATCCAG 906
 QY 62 CTTCCCGTTATGAAGTCGCAACGTGTCGCGGGGTACCATGTACCAACGACTGCTCCA 121
 DB 907 CTTCCCGTTATGAAGTCGCAACGTGTCGCGGGATATACCATGTACCAACGACTGCTCCA 966
 QY 122 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGCGCCT 181
 DB 967 ACTCAAGCATAGTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGCGCCT 1026
 QY 182 GCGTTCCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACCGTCGCGGCCA 241
 DB 1027 GTGTTTCAGAGGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACCGTCGCGGCCA 1086
 QY 242 GGAAGCGCAGCGTCCCAACGACATACGACGCGACGTCGATTCGCTGTTGGGCTG 301
 DB 1087 GGAATGCCAGCGTCCCAACGACATACGACGCGACGTCGATTCGCTGTTGGGACGG 1146
 QY 302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTCTGTTTCCC 361
 DB 1147 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGGATCTATTTTCTGCTCCC 1206
 QY 362 AGCTGTTCACTTCTCACTCGCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
 DB 1207 AGCTGTTCACTTCTCGCTCGCGCGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
 QY 422 ATCCCGCCCATGTATCAGGTCACCGCATGCGCTTGGGATATGATGAACTGGTC 476
 DB 1267 ATCCCGCCCATGTATCAGGTCACCGCATGCGCTTGGGATATGATGAACTGGTC 1321

RESULT 14

AA24833
 ID AAX24833 standard; DNA; 9599 BP.

XX
 AC AAX24833;

XX
 DT 21-JUN-1999 (first entry)

XX
 DE Infectious hepatitis C virus genotype 1a/1b chimera genome.

XX
 KW HCV; infectious clone; infection; diagnosis; therapy; vaccine;
 KW screening; assay; antiviral; virucide; ss.

XX
 OS Hepatitis C virus.

XX
 FH Key Location/Qualifiers

XX
 CDS 342..9377

XX
 FT /*tag= a

XX
 PN WO9904008-A2.

PD XX 28-JAN-1999.

PX PF 16-JUL-1998; 98WO-US14688..

PR PR 27-JAN-1998; 98US-0014416.

XX PA 18-JUL-1997; 97US-0053062.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

PI BUkh J, Emerson SU, Purcell RH, Yanagi M;

DR WIPI; 1999-132252/11.

DR P-PADB; AAW98021.

XX New isolated hepatitis C virus nucleic acids - used to develop products for the diagnosis, prevention and treatment of HCV infections and for developing screening assays

PT Claim 8; Fig 16A-F; 126pp; English.

PS This nucleic acid sequence comprises the genome of infectious hepatitis C virus (HCV) chimeric la/lb clone pH7C-J4 which contains the nonstructural region of infectious genotype 1a strain H77 (see AAX24832) and the structural region of infectious genotype lb strain HC-J4 (see AAX24843). The construction of such chimERIC nucleic acid sequences is expected to be of importance in studying the growth and virulence properties of HCV and in the production of HCV suitable for conferring protection against multiple genotypes of HCV. The invention also relates to the introduction of mutations or deletions into infectious nucleic acid sequences in order to produce an attenuated HCV virus suitable for vaccine development. Infectious nucleic acid sequences can also be used to produce attenuated virus via passage in vitro or in vivo of the viruses produced by transfection of a host cell with the infectious nucleic acid sequence. Vaccines comprising one or more polypeptides made from the infectious nucleic acid sequence are used to immunise mammals, especially humans, against hepatitis C. The nucleic acid sequences can also be used to induce protective immunity against the virus. The nucleic acid sequences or their encoded proteases (e.g. NS3 protease), can additionally be used to develop screening assays to identify antiviral agents for HCV.

SQ Sequence 9599 BP; 1889 A; 2873 C; 2724 G; 2113 T; 0 other;
Query Match 89.1%; Score 430.2; DB 20; Length 9599;
Best Local Similarity 94.1%; Pred.No. 2.1e-115;
Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2 TGCCCGGTGGCTTTTCTATTCCTTGCGGCTGCCTGCTGTAACCATACCAG 61
DB 847 TGCCCGGTGGCTTTTCTATTCCTTGCGGCTGCCTGCTGTAACCATACCAG 906
QY 62 TTTCGCCATTGAAGTGCACAACGTGTCGGGGTGTAACCAATGACAAAGACTGTCCA 121
DB 907 TTTCGCCATTGAAGTGCACAACGTGTCGGGGTAGAACATAACCAATGACAAAGACTGTCCA 966
QY 122 ACTCAAGCATAGTGTATGAGCGCGGATGATCATGSCACACCCC CGGGTGCGTGCCT 181
DB 967 ACTCAAGCATAGTGTATGAGCGCGGATGATCATGSCATCTACCTCCCGGGTGCGTGCCT 1026
QY 182 GGTTTTCGGGAGGCGCAACTCTCCCTTGCTGGGTGGGCTCACATCCACCGTTCGGGCGCA 241
DB 1027 GTGTTCAGAGGTTAACAGCTCCCGTGTCTGGGTAGCGCTCACTCCCCACCGTTCGGGCGCA 1086
QY 242 GGAAAGCCAGCGTCCCGCAACGCAATACGACCGGCGCTCGATTTGCTTGGGGCTG 301
DB 1087 GGAATGCCAGCGTCCCGCACTACGACAAATACGACCGGCGCTCGATTTGCTTGGGGCTG 1146
QY 302 CTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTCTGGGGATCTCTTTTCTTTTCTTCCC 361
DB 1147 CTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTCTGGGGATCTATTTTCTTCTCTCCC 1206
QY 362 AGCTGTTCACTTTCTCACCTCGCCGCGCATCAAACAGTACGAGCTGCAACTGCTCAATCT 421

Qy	302	CTGCTTTCTGTCGGCTATGTACGTGGGGATCTCTGGGATCTGTGTTTCTGTTTCCC	361
Db	362	CTGCTTCTGTCCTCGCTATGTACGTGGGAGATCTCTGGGATCTGTGTTTCTGTTTCCC	421
Qy	362	AGCTGTTCACTTCTCACCTTCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATCT	421
Db	422	AGCTGTTCACTTCTCACCTTCGCGGCATGAGACAGTACAGGACTGCAATTGCTCAATCT	481
Qy	422	ATCCGGCCATGTATCAGGTACCGCATGGCTTGGGATATGATGATGAATGGTTC	476
Db	482	ATCTGGCCACGTATCAGGTATCGCATGGCTTGGGATATGATGATGAATGGTTC	536

Search completed: December 19, 2003, 18:51:05
Job time : 135.529 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1242.55 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATGCCGGTGTCTTCTC.....TGATGACTGGTCCTAATAG 483

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_man:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.2	13.9	488	9	AV755731
C 2	55.6	11.5	492	9	AV758366
C 3	43.2	8.9	534	14	CD040840
C 4	41.6	8.6	664	29	BZ645446

AV755731 AV755731
AV758366 AV758366
CD040840 p9HB036XB
BZ645446 OGCB086TC

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguaguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
FEATURES

C 5 40.6 8.4 526 9 AL825643
6 40.2 8.3 399 9 AV638521
7 40.2 8.3 434 9 AV637507
8 40.2 8.3 440 9 AV637983
9 40.2 8.3 450 9 AV637259
10 40.2 8.3 451 9 AV637328
11 40.2 8.3 451 9 AV637643
12 40.2 8.3 453 9 AV634724
13 40.2 8.3 454 9 AV637050
14 40.2 8.3 456 9 AV635382
15 40.2 8.3 473 9 AV632765
16 40.2 8.3 481 9 AV635503
17 40.2 8.3 485 9 AV632811
18 40.2 8.3 506 9 AV392445
19 40.2 8.3 508 9 AV634095
20 40.2 8.3 526 9 AV641895
21 40.2 8.3 533 9 AV638125
22 40.2 8.3 537 9 AV632335
23 40.2 8.3 588 9 AV387329
24 40.2 8.3 1186 13 BX421743
25 39.4 8.2 1201 9 AL565958
C 26 39.2 8.1 624 14 CD206870
C 27 39.2 8.1 656 14 CB924688
C 28 39 8.1 497 9 AV633658
29 39 8.1 610 14 CB657655
30 39 8.1 856 29 BZ578381
31 39 8.1 872 29 BZ555011
C 32 38.4 8.0 645 29 CNS01213
C 33 38 7.9 771 29 BZ530934
C 34 37.8 7.8 490 9 AV634529
C 35 37.8 7.8 705 14 CA618797
C 36 37.8 7.8 1039 13 BX415186
C 37 37.6 7.8 309 12 BI098866
C 38 37.6 7.8 431 9 AV639153
C 39 37.4 7.7 394 12 BJ209795
C 40 37.4 7.7 501 9 AV638474
C 41 37.2 7.7 431 9 AV641448
C 42 37.2 7.7 792 13 BX391120
C 43 37 7.7 910 29 CNS0060N
44 36.8 7.6 214 12 BM686105
45 36.8 7.6 913 14 CA487901

AL825643 AL825643
AV638521 AV638521
AV637507 AV637507
AV637983 AV637983
AV637259 AV637259
AV637328 AV637328
AV637643 AV637643
AV634724 AV634724
AV637050 AV637050
AV635382 AV635382
AV632765 AV632765
AV635503 AV635503
AV632811 AV632811
AV392445 AV392445
AV634095 AV634095
AV641895 AV641895
AV638125 AV638125
AV632335 AV632335
AV387329 AV387329
BX421743 BX421743
AL565958 AL565958
CD206870 HSI_25_D0
CB924688 ABAL_1_D1
AV633658 AV633658
CB657655 OSJNBL13C
BZ578381 msh2_5817
BZ555011 pacsl-60
AL101589 Drosophila
BZ530934 OGAVJ23TC
AV634529 AV634529
CA618797 wlin.pk00
BX415186 BX415186
BI098866 IPL_34_CO
AV639153 AV639153
BJ209795 BJ209795
AV638474 AV638474
AV641448 AV641448
BX391120 BX391120
AL065629 Drosophila
BM686105 mgcs013xi
CA487901 AGENCOURT

```
source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT      116 a 134 c 137 g 97 t      4 others
ORIGIN
Query Match      13.9%; Score 67.2; DB 9; Length 488;
Best Local Similarity 67.0%; Pred. No. 1.1e-07;
Matches 126; Conservative 0; Mismatches 58; Indels 4; Gaps 2;

QY 292 GTTGGGGCTGCTTCTTCTGTTCTGCTATGCTAGTGGGGATCTCTGCGGATCTGCGGATCTGTTTC 351
Db 472 GTGGTGCACACTCGCTCTGCTCAGCTCTCTAGTGTGGGACCTCTGCGGAGGTG 413
QY 352 GTTGTTCACAGTGTTCACCTTCTCAGTGGCGGCATCAACAGTACAGGACTGCAAC 411
Db 412 CTTGAGTTCAGTGATCA---TCTGGCCTCAGCACCATGAGTTGTGCATGATGCAAC 356
QY 412 TGCTCATCTATCCGGGCATGTATCAGTCAACGCATG-GCTTGGGATATGATGAA 470
Db 355 TGCTCATCTATCTGGGGCCATCACTGGACACCGTATGACGATGGGACATGATGAA 296
QY 471 CTGTCCT 478
Db 295 CTGTCGT 288

RESULT 2
AV758366/c
LOCUS      AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA linear EST 19-OCT-2000
DEFINITION      AV758366
ACCESSION      AV758366
VERSION      AV758366.1 GI:10916214
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 492)
AUTHORS      Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng
, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE      Homo sapiens cDNA BM clones
JOURNAL      Unpublished
COMMENT      Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT      124 a 128 c 125 g 112 t      3 others
ORIGIN

Query Match      11.5%; Score 55.6; DB 9; Length 492;
Best Local Similarity 65.9%; Pred. No. 0.00016;
Matches 112; Conservative 0; Mismatches 54; Indels 4; Gaps 2;

QY 308 TCTGTTCCGCTATGATACGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCAGTCT 367
Db 457 TGTGATCAGTCTACTACGTGTGGACCTCTGCGTTGGGGTATCGCTTGCAGCCCACTGA 398
QY 368 TCACCTTCTCACCTGCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCTATCCCG 427
Db 397 TTA---TCTCTCAGCAGCAACATTTGGTTTGTGCAAGATGCAACTGCTCATTTCTATCTG 341
QY 428 GCCATGTATCAGGTC-ACCGCATGGCTTGGGATATGATGATGAACCTGGTC 476
Db 340 GCTGCATCACTGGACTACAGTATGGCATAGGCTATGATGATGAACCTGGTC 291

RESULT 3
CD040840      534 bp mRNA linear EST 09-MAY-2003
psHB036x809f_300663 psHB: Infected hypocotyl soybean host. 48 hrs
post infection Phytophthora sojae cDNA clone SHB036B09 5', mRNA
sequence.
ACCESSION      CD040840
VERSION      CD040840.1 GI:30502701
KEYWORDS      EST.
SOURCE      Phytophthora sojae
ORGANISM      Phytophthora sojae
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.
REFERENCE      1 (bases 1 to 534)
AUTHORS      Tyler, B.M., Judelson, H.S., Gijzen, M., Dean, R.A. and Waugh, M.E.
TITLE      USDA-IPAFS: Expression of Phytophthora sojae genes during infection
and propagation
JOURNAL      Unpublished
COMMENT      Contact: Tyler B
Tyler lab
VBI
1880 Pratt Dr., Blacksburg, VA 24061, USA
Tel: 540-231-7318
Email: bmtyle@vt.edu
PCR Primers
FORWARD: BK reverse
Plate: 036 row: B column: 09
Seq primer: BK reverse
High quality sequence stop: 534.
FEATURES
source
1..534
/organism="Phytophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="aHB036B09"
/tissue_type="infected host tissue"
/cell_line="P6497"
/dev_stage="48 hour post infection"
/clone_lib="psHB: Infected hypocotyl soybean host. 48 hrs
post infection"
/notes="vector: psBK-CMV, Site_1: EcoRI, Site_2: XhoI;
USDA-IPAFS: Expression of Phytophthora sojae genes during
infection and propagation."
BASE COUNT      101 a 187 c 159 g 87 t
ORIGIN

Query Match      8.9%; Score 43.2; DB 14; Length 534;
Best Local Similarity 47.4%; Pred. No. 0.39; 143; Indels 0; Gaps 0;
Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGCACAGTCTCGGGGTGTACCATCTCAGCAACGACTGCTGCTCAACTCAGC 129
Db 200 TAGCGCGTCCGGAGATTACGGTATCGCATGGCTTTCGCCGCTTCTACAACTGGACC 259
QY 130 ATAGTGTATGAGGCGAGCATGATCATGCACACCCCGGGTGGTGGCTTGGCTTGG 189
```


Mon Dec 22 13:28:55 2003

```

adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
    source
    1..399
    /organism="Chlamydomonas reinhardtii"
    /mol_type="mRNA"
    /db_xref="taxon:3055"
    /clone_lib="HC074a01_r"
    /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
    xhoI; The cDNA library was constructed from cells cultured
    in a medium with bubbling air containing 5% carbon
    dioxide"
    70 a 144 c 122 g 63 t
    BASE COUNT
    ORIGIN
    Query Match 8.3%; Score 40.2; DB 9; Length 399;
    Best Local Similarity 51.4%; Pred. No. 2.3; Mismatches 0; Gaps 0;
    Matches 93; Conservative 0; Indels 88; Indels 0; Gaps 0;

    QY 30 CTTGGCCCTGTCCTGTCGACATACCAAGCTTCGCTTATGAAGTGGCAACGTGTC 89
    Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    44 CTTGACCCCTGGACGGCTCGTCATCGTCGACCAACTTCGCCGTGGCACCATTCTT 103
    QY 90 CGGGGTGTACCATGTCACGAACGACTGCTCCAACTCAAGCATAGTGTATGAGGAGCGGA 149
    Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    104 CGCGGTGTGGCCGGCGACAGCTGACCAACATCACCGCGCGAGCAGGTGGTGGCGG 163
    QY 150 CATGATCATGCACACCCCGGGTGGCTGCGCTTGGGAGGCAACTCTCCCGTTG 209
    Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    164 CATGGCATCTACGGTCCCGCAGCGGTGTTCTGCAATTCCTGAAGGACGCCCGCGCTG 223
    QY 210 C 210
    Db 224 C 224

    RESULT 7
    AV637507 434 bp mRNA linear EST 15-DEC-2000
    LOCUS AV637507 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
    DEFINITION cDNA clone HC074a01_r 5', mRNA sequence.
    ACCESSION AV637507
    VERSION AV637507.1 GI:10780827
    KEYWORDS EST.
    SOURCE Chlamydomonas reinhardtii
    ORGANISM Chlamydomonas reinhardtii
    Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
    Chlamydomonadaceae; Chlamydomonas.
    REFERENCE 1 (bases 1 to 434)
    AUTHORS Asamizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
    Nakamura,Y. and Tabata,S.
    COMMENT Generation of expressed sequence tags from low-CO2 and high-CO2
    adapted cells of Chlamydomonas reinhardtii
    DNA Res. 7 (5), 305-307 (2000)
    PUBMED 20539644
    CONTACT: Erika Asamizu
    The First Laboratory for Plant Gene Research
    Kazusa DNA Research Institute
    Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
    Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
    Location/Qualifiers
    1..434
    /organism="Chlamydomonas reinhardtii"

```

```

/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone_lib="HC074a01_r"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
xhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
80 a 149 c 130 g 75 t
BASE COUNT
ORIGIN
    Query Match 8.3%; Score 40.2; DB 9; Length 434;
    Best Local Similarity 51.4%; Pred. No. 2.4; Mismatches 0; Gaps 0;
    Matches 93; Conservative 0; Indels 88; Indels 0; Gaps 0;

    QY 30 CTTGGCCCTGTCCTGTCGACATACCAAGCTTCGCTTATGAAGTGGCAACGTGTC 89
    Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    21 CTTGACCCCTGGACGGCTCGTCATCGTCGACCAACTTCGCCGTGGCACCATTCTT 80
    QY 90 CGGGGTGTACCATGTCACGAACGACTGCTCCAACTCAAGCATAGTGTATGAGGAGCGGA 149
    Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    81 CGCGGTGTGGCCGGCGACAGCTGACCAACATCACCGCGCGAGCAGGTGGTGGCGG 140
    QY 150 CATGATCATGCACACCCCGGGTGGCTGCGCTTGGGAGGCAACTCTCCCGTTG 209
    Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    141 CATGGCATCTACGGTCCCGCAGCGGTGTTCTGCAATTCCTGAAGGACGCCCGCGCTG 200
    QY 210 C 210
    Db 201 C 201

    RESULT 8
    AV637983 440 bp mRNA linear EST 15-DEC-2000
    LOCUS AV637983 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
    DEFINITION cDNA clone HC080c04_r 5', mRNA sequence.
    ACCESSION AV637983
    VERSION AV637983.1 GI:10781303
    KEYWORDS EST.
    SOURCE Chlamydomonas reinhardtii
    ORGANISM Chlamydomonas reinhardtii
    Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
    Chlamydomonadaceae; Chlamydomonas.
    REFERENCE 1 (bases 1 to 440)
    AUTHORS Asamizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
    Nakamura,Y. and Tabata,S.
    COMMENT Generation of expressed sequence tags from low-CO2 and high-CO2
    adapted cells of Chlamydomonas reinhardtii
    DNA Res. 7 (5), 305-307 (2000)
    PUBMED 20539644
    CONTACT: Erika Asamizu
    The First Laboratory for Plant Gene Research
    Kazusa DNA Research Institute
    Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
    Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
    Location/Qualifiers
    1..440
    /organism="Chlamydomonas reinhardtii"
    /mol_type="mRNA"
    /strain="C9"
    /db_xref="taxon:3055"
    /clone_lib="HC080c04_r"
    /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
    xhoI; The cDNA library was constructed from cells cultured
    in a medium with bubbling air containing 5% carbon
    dioxide"
    76 a 156 c 137 g 71 t
    BASE COUNT
    ORIGIN

```



```

/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC068a04_1"
/clone_lib="Chlamydomonas reinhardtii 5' CO2"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
82 a 153 c 144 g 75 t
BASE COUNT
ORIGIN

```

Query Match	8.3%;	Score 40.2;	DB 9;	Length 454;
Best Local Similarity	51.4%;	Pred. No. 2.4;		
Matches	93;	Conservative 0;	Mismatches 88;	Indels 0;
Gaps	0;			
Qy	30	CTTGGCCCTGCTGCTCTGTCATGACATACAGCTTCGGCTTATGAAGTCGGCAACGTGC	89	
Db	52	CTTCGACCCCTGGACGGCTCGTCCATCGTCGACACCAACTTCGCCGTGGCACCATCTT	111	
Qy	90	CGGGGTGTACATGTCTACGAAACGACTGCTCCAACTCAAGCATATGCTPATGAGCAGCGGA	149	
Db	112	CGGGGTGTGGCCCGGCACAAAGCTGACCAACATCACCGCCGGCAGCAGTGGCTGCCGG	171	
Qy	150	CATGATCATGCACACCCCGGGTCGGTCCTCGGTCGGGAGGGCAACTCTCCCGTTG	209	
Db	172	CATGGGCATCTACGGTCCCCGCACCGGTGTCTGTGATTCGCCCTGAAGACGCCCCCGGCTG	231	
Qy	210	C 210		
Db	232	C 232		

RESULT 14	
AV635382	
LOCUS	456 bp mRNA linear EST 15-DEC-2000
DEFINITION	AV635382.Chlamydomonas reinhardtii 5' CO2 Chlamydomonas reinhardtii cDNA clone HC045f10_r 5', mRNA sequence.
ACCESSION	AV635382
VERSION	AV635382.1 GI:10778702
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE	1 (bases 1 to 456)
AUTHORS	Asamizu,Y., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,E. and Tabata,S.
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
JOURNAL	DNA Res. 7 (5), 305-307 (2000)
MEDLINE	20539644
PUBMED	11089912
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/ . Location/Qualifiers 1..456
FEATURES	
SOURCE	

```

1. 456
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC045f10_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="vector: BluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
75 a 158 c 148 g 75 t
BASE COUNT
ORIGIN

```

Query Match	8.3%	Score 40.2	DB 9	Length 456
Best Local Similarity	51.4%	Pred. No. 2.4		
Matches 93	Conservative 0	Mismatches 88	Indels 0	Gaps 0
Qy	30	CTTGGCCCTGCTGTCGTGTGACCATACAGCTTCGGCTATATGAAGTCGGCAACGTC	89	
Db	221	CTTCGACCCCTGGACGGCTGTCATCTGTCGACACCAACTTCGCCGTGGCACCATCTT	280	
Qy	90	CGGGGTGTACATGTCTACGAACGACTGCTCCAACTCAAGCATATAGTATAGGAGCGCGGA	149	
Db	281	CGGCGTGTGGCCGGCGACAGCTGACCAACATACCGCCGGCAGCAGGTGGCTGCCGG	340	
Qy	150	CATGATCATGCACACCCCGGGTCGGTCCCTCGTTTCGGAGGGCAACTCTCCCGTTG	209	
Db	341	CATGGGCATCTAGGTCCTCCCGACCGTGTTCTGCATTGCCCTGAAGGACGCCCCGCGCTG	400	
Qy	210	C 210		
Db	401	C 401		

RESULT 15	AV632765	473 bp	mRNA	linear	EST 15-DEC-2000
LOCUS	AV632765				
DEFINITION	Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii cDNA clone HC012c10_r 5', mRNA sequence.				
ACCESSION	AV632765				
VERSION	AV632765.1				
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii				
ORGANISM	Chlamydomonas reinhardtii				

REFERENCE	Chlamydomonadales; Chlamydomonas.
AUTHORS	1 (bases 1 to 473) Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y., Miura, K. and Tabata, S.
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
JOURNAL	DNA Res. 7 (5), 305-307 (2000)
MEDLINE	20539644
PUBMED	11089912
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/ .
FEATURES	Location/Qualifiers 1..473

```

1. 473
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC012c10_r"
/clone_lib="Chlamydomonas reinhardtii 5' CO2"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
82 a 160 c 151 g :80 t
BASE COUNT
ORIGIN

```

	Query Match	8.3%	Score 40.2;	DB 9;	Length 473;	
	Best Local Similarity	51.4%;	Pred. No. 2.5;			
	Matches	93;	Conservative	0; Mismatches	88; Indels	0; Gaps
Qy	30	CTTGGCCCTGCTGTCTTGACCATPACAGCTTCGGCTTATGAAGTGCCEAACGTGTC	89			
Dd	156	CTTCGACCCCTGGACGGCTCGTCCATCGTCGACACCACCACTTCGCGTGGGCCACCATCTT	215			
Qy	.90	CGGGGTGTTACATGTCTACAGAACGACTGCTCCAACCTAAGCATATGATGTATGAGGACGGGA	149			
Dd	216	CGGCGTGTGGCCCGGGCACAAGCTGACCAACATCACCGGCGGGCAGAGTGGCTGCGCG	275			

Mon Dec 22 13:28:55 2003

```
QY      150 CATGATCATGCACACCCCGGGTGGTGCCCTGCGTTGGGAGGGGAACTCCTCCCGTTG 209
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
          276 CATGGGCATCTACGGTCCCGGACACCGTGTTCAGCATTCCTGAAGGACGCCCCCGGCTG 335
QY      210 C 210
Db      336 C 336
```

Search completed: December 20, 2003, 06:54:43
Job time : 1246.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 34.8338 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATCCCGGTGCTCTTCTC.....TGATGAACGTGCTCTAATAG 483

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483	100.0	483	3	US-08-612-973-9
2	483	100.0	483	3	US-08-927-597-9
3	474.2	98.2	480	3	US-08-612-973-11
4	474.2	98.2	480	3	US-08-927-597-11
5	433.4	89.7	1539	2	US-08-470-426B-17
6	433.4	89.7	1863	2	US-08-470-426B-14
7	430.2	89.1	9595	3	US-09-014-416-4
8	430.2	89.1	9599	3	US-09-014-416-6
9	425.4	88.1	9472	4	US-08-150-204E-96
10	425	88.0	501	2	US-08-483-695-30
11	425	88.0	501	2	US-07-965-285-30
12	425	88.0	501	2	US-08-487-231-30
13	425	88.0	501	3	US-09-201-912-30
14	424.8	88.0	642	3	US-08-612-973-3
15	424.8	88.0	642	3	US-08-927-597-3
16	423.8	87.7	795	3	US-08-612-973-5
17	423.8	87.7	795	3	US-08-927-597-5
18	423.8	87.7	2082	3	US-08-612-973-47
19	423.8	87.7	2082	3	US-08-927-597-47
20	423.8	87.7	2433	3	US-08-612-973-49
21	423.8	87.7	2433	3	US-08-927-597-49
22	423.6	87.7	633	3	US-08-612-973-7
23	423.6	87.7	633	3	US-08-927-597-7
24	422.8	87.5	636	3	US-08-612-973-13
25	422.8	87.5	636	3	US-08-927-597-13
26	422.2	87.4	1037	1	US-08-462-195-1
27	422.2	87.4	1037	2	US-08-636-883-1

28 422.2 87.4 1037 3 US-09-127-829-1 Sequence 1, Appli

29 419 86.7 742 1 US-08-081-072-18 Sequence 18, Appl

30 419 86.7 742 1 US-08-449-093A-18 Sequence 18, Appl

31 419 86.7 932 1 US-08-081-072-15 Sequence 15, Appl

32 419 86.7 932 1 US-08-449-093A-15 Sequence 15, Appl

33 417.4 86.4 2116 3 US-08-191-160-21 Sequence 21, Appl

34 415.4 86.0 501 2 US-08-483-695-28 Sequence 28, Appl

35 415.4 86.0 501 2 US-07-965-285-28 Sequence 28, Appl

36 415.4 86.0 501 3 US-08-487-231-28 Sequence 28, Appl

37 415.4 86.0 501 3 US-09-201-912-28 Sequence 28, Appl

38 407.8 84.4 1167 1 US-08-324-977-9 Sequence 9, Appli

39 407.8 84.4 1167 2 US-08-384-616-9 Sequence 9, Appli

40 407.8 84.4 1167 2 US-08-904-686A-9 Sequence 9, Appli

41 407.8 84.4 1499 1 US-09-315-850-9 Sequence 3, Appli

42 407.8 84.4 1499 2 US-08-324-977-3 Sequence 3, Appli

43 407.8 84.4 1499 2 US-08-384-616-3 Sequence 3, Appli

44 407.8 84.4 1499 2 US-08-904-686A-3 Sequence 3, Appli

45 407.8 84.4 1499 3 US-09-315-850-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-08-612-973-9

; Sequence 9, Application us/08612973

; Patent No. 6150134

; GENERAL INFORMATION:

; APPLICANT: MAERTENS, GEERT

; APPLICANT: BOSMAN, FONS

; APPLICANT: DE MARTYNOFF, GUY

; APPLICANT: BUYSSE, MARIE-ANGE

; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

; NUMBER OF SEQUENCES: 111

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/612.973

FILING DATE: 11-MAR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 483 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..480

FEATURE:

NAME/KEY: mat.peptide

LOCATION: 1..477

Mon Dec 22 13:28:54 2003

```

Db      61  |||||GCTTCGCGTTATGAAGTGGCGCAACGTTGTCGGGGGTGTACATGTCAAGAACGACTGCTCC 120
Qy      121  AACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGCACACCCCGGGTGGCTGCC 180
Db      121  AACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGCACACCCCGGGTGGCTGCC 180
Qy      181  TGGCTTCGGAGGGCAACTCTCCGTTGCTGGGTGGCGTCACTCCACGCTCGCGGCC 240
Db      181  TGGCTTCGGAGGGCAACTCTCCGTTGCTGGGTGGCGTCACTCCACGCTCGCGGCC 240
Qy      241  AGGAACCGCAGCGTCCCAACAACGACAATACGACGCGCACGTCGAMTTGCTGTGGGGCT 300
Db      241  AGGAACCGCAGCGTCCCAACAACGACAATACGACGCGCACGTCGAMTTGCTGTGGGGCT 300
Qy      301  GCTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTGCGGATCTGTTTTCCTTTGTTCC 360
Db      301  GCTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTGCGGATCTGTTTTCCTTTGTTCC 360
Qy      361  CAGCTGTTCACCTTCTCACTTCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATC 420
Db      361  CAGCTGTTCACCTTCTCACTTCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATC 420
Qy      421  TATCCCGGCCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAACTGGTAATA 479
Db      421  TATCCCGGCCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAACTGGTAATA 479

RESULT 5
US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

```

```

US-08-470-426B-17
Query Match      89.7%; Score 433.4; DB 2; Length 1539;
Best Local Similarity 94.5%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
Matches 449; Conservative 0; Mismatches 26;

Qy      2  TGGCCGGTTCCTTTCTCTATCTTCTTGGCCCTGCTGCTCTGTGACCATACCAG 61
Db      506  TGGCCGGTTCCTTTCTCTATCTTCTTGGCCCTGCTGCTCTGTGACCATACCAG 565
Qy      62  CTTCCGGTTATGAAGTGGCAACGTTGTCGGGGGTGTACCATGTCCAGAACGACTGCTCCA 121
Db      566  CTTCCGGTTATGAAGTGGCAACGTTGTCGGGGGTGTACCATGTCCAGAACGACTGCTCCA 625
Qy      122  ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGCACACCCCGGGTGGCTGCCCT 181
Db      626  ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGCACACCCCGGGTGGCTGCCCT 685
Qy      182  GCGTTTCGGGAGGGCAACTCTCCGTTGCTGGGGGTGCTCACTCCACGCTCGCGGCCA 241
Db      686  GCGTTTCGGGAGGCAACTCTCCGTTGCTGGGGGTGCTCACTCCACGCTCGCGGCCA 745
Qy      242  GGAACCGCAGCGTCCCAACAACGACAATACGACGCGCACGTCGAMTTGCTGTGGGGCTG 301
Db      746  GGAACCGCAGCGTCCCAACAACGACAATACGACGCGCACGTCGAMTTGCTGTGGGGCTG 805
Qy      302  CTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTCTGCGGATCTGTTTTCCTTTGTTCC 361
Db      806  CTGCTTTCTGTTCCGCTATGATAGTGGGGGATCTCTGCGGATCTGTTTTCCTTTGTTCC 865
Qy      362  AGCTGTTCACCTTCTCACTTCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATCT 421
Db      866  AGCTGTTCACCTTCTCACTTCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATCT 925
Qy      422  ATCCCGGCCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAACTGGTTC 476
Db      926  ATCCCGGCCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAACTGGTTC 980

RESULT 6
US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

```

```

; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1863 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-470-426B-14

Query Match      89.7%; Score 433.4; DB 2; Length 1863;
Best Local Similarity 94.5%; Pred. No. 1.4e-112;
Matches 449; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY    2   TGCCCGGTTGCCTTTCCTATCTCCTCTTGCCCTCTGTCTCTGCTGACCATACCAG 61
DB    830 TGCCCGGTTGCCTTTCCTATCTCCTCTTGCCCTCTGTCTCTGCTGACCATACCAG 889

QY    62   CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTGCACGAACGACTGCTCCA 121
DB    890 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTGCACGAACGACTGCTCCA 949

QY    122 ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGCACACCCCGGGTGCCTGCCCT 181
DB    950 ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGCACACCCCGGGTGCCTGCCCT 1009

QY    182 GGGTTCCGGAGGCGCAACCTCCTCCCTGCTGGGTGGCGCTCACCTCCACGCTCGCGGCCA 241
DB    1010 GGGTTCCGGAGGCGCAACCTCCTCCCTGCTGGGTGGCGCTCACCTCCACGCTCGCGGCCA 1069

QY    242 GGAAAGCCACGCTCCCCAACACAACAGCAACATACGACGCGCTCGATTGCTCTGTTGGGGCTG 301
DB    1070 GGAAAGCCAGGCTCCCACATACGACCAATACGAGCGCACTGCTGCTGGGGGG 1129

QY    302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGCGATCTGTTTTCTGTTTCCC 361
DB    1130 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGCGATCTGTTTTCTGCTCTCCC 1189

QY    362 AGCTGTTCACTTCTCACTCGCCGCGCATCAAACAGTACAGACTGCAACTGCTCAATCT 421
DB    1190 AGCTGTTCACTTCTCACTCGCCGCGCATGACAGAGTGCAGGACTGCAACTGCTCAATCT 1249

QY    422 ATCCCGGCGCATGTATCAGGTACCGCATCGCTTGGGATATGATGATGAACCTGGTC 476
DB    1250 ATCCCGGCGCATGTATCAGGTACCGCATCGCTTGGGATATGATGATGAACCTGGTC 1304

RESULT 7
US-09-014-416-4
; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014, 416
; EARLIER FILING DATE: 1998-01-27
; EARLIER FILING DATE: 1997-07-18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 9595
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-014-416-4

Query Match      89.1%; Score 430.2; DB 3; Length 9595;
Best Local Similarity 94.1%; Pred. No. 1.8e-111;
Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY    2   TGCCCGGTTGCCTTTCCTATCTCCTCTTGCCCTCTGTCTCTGCTGACCATACCAG 61
DB    847 TGCCCGGTTGCCTTTCCTATCTCCTCTTGCCCTCTGTCTCTGCTGACCATACCAG 906

QY    62   CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTGCACGAACGACTGCTCCA 121
DB    907 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTGCACGAACGACTGCTCCA 966

Query Match      89.1%; Score 430.2; DB 3; Length 9599;
Best Local Similarity 94.1%; Pred. No. 1.8e-111;
Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY    2   TGCCCGGTTGCCTTTCCTATCTCCTCTTGCCCTCTGTCTCTGCTGACCATACCAG 61
DB    847 TGCCCGGTTGCCTTTCCTATCTCCTCTTGCCCTCTGTCTCTGCTGACCATACCAG 906

QY    62   CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTGCACGAACGACTGCTCCA 121
DB    907 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGTGTACCATGTGCACGAACGACTGCTCCA 966

```


QY	242	GGAAAGCCAGCGTCCCAACAAGCAAAATACGACGCCACGTCGATTTGCTCGTTGGGGCTG	301
Db	277	GGNATGCGACGCTCCCACTACGACAAATACGACGCCACGTCGATTTGCTCGTTGGGGCGG	336
QY	302	CTGCTTTCTGTTCCGCTATGTAGTGGGGGATCTCTGCGGATCTGTTTTCCTTGTTTCCC	361
Db	337	CTGCTTTCTGCTCCGCTATGTAGTGGGGGATCTCTGCGGATCTGTTTTCCTGCTCCC	396
QY	362	AGCTGTTTCACCTTCTCACCTCGCCGGCATCAAACAGTACAGGACTGGCACTGCTCAATCT	421
Db	397	AGCTGTTTCACTTCTCGCCTCGCGGCATGAGACAGTGCAGGACTGCAACTGCTCAATCT	456
QY	422	ATCCCGGCCATGTATCAGGTACCGCATGGCTTTGGGATATGATGA	466
Db	457	ATCCCGGCCATTTATACAGGTACCGCATGGCTTTGGGATATGATGA	501

```

RESULT 13
US-09-201-912-30
; Sequence 30, Application US/09201912
; Patent No. 6210962
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremsdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; CURRENT APP. NUMBER: US/09/201-912

```

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,285
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
DESCRIPTION: cDNA to genomic RNA
US-09-201-912-30

```

Query Match	88.0%	Score 425;	DB 3;	Length 501;
Best Local Similarity	94.6%	Pred. No. 2.3e-110;		
Matches 440; Conservative	0;	Mismatches 25;	Indels 0;	Gaps 0;
QY	2	TGCCCGGTGCTTTCTCTATCTTCTCTTGGCCCTGCTGTCCTGCTGACCATACCAG	61	

```

RESULT 12
US-08-487-231-30
; Sequence 30, Application US/08487231
; Patent No. 5919454
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Finnegan, Henderson, Farabow, Garrett &
;

```

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,231
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,285
FILING DATE: 18-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 06 882
FILING DATE: 06-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000

```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 501 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: Other
;   DESCRIPTION: CDNA to genomic RNA
US-08-487-231-30

Query Match      88.0%; Score 425; DB 2; Length 501;
Best Local Similarity 94.6%; Pred. No. 2.3e-110;
Matches 440; Conservative 0; Mismatches 25; Indels 0; Gaps 0

QY      2  TGCCCGGTGCTCTTCTATCTCTTCCTTGTCGCCCTGCTGTCCTGTCTGACCATACCAG 61
Db      37  TGCCCGGTGCTCTTCTCTATCTCTTCCTTGCTGCCGTGCTGTCCTGTTGACCATCCAG 96

QY      62  CTTTCGCCTTATGAAGTGCCCAACGTTGTCGGGGTGTAACATGTCACGAACGACTGCTCCA 121
Db      97  CTTTCGCCTTATGAAGTGCCCAACGTTGTCGGGATATACCATGTCACGAACGACTGCTCCA 156

```

122	ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGACACCCCGGGTCGTGCCCT	181
QY		
Db		
157	ACTCAAGCATGTGTATGAGGAGCGGACATGATCATGATGCTACTCCGGGTGGTGCCT	216
QY		
182	GGTTCGGGAGGSCAATCCTCCGTTGTTGGGTGGCGCTCACTCCACCGTTCGGGCCA	241
QY		

37	TGCCGGGTGCTCTTTCTCTATCTCTCTTGGCTTTGCTGCTGCTGTTGACCATCCCA	96
62	CTTCCGCTTATGAAGTGGCAACGTTGTCGGGGTGTACCATGTCAAGAACGACTGCTCCA	121
97	CTTCCGCTTATGAAGTGGCAACGTTGTCGGGGATATACCATGTCAAGAACGACTGCTCCA	156
122	ACTCAAGCATAGTGTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGTGCCCT	181
157	ACTCAAGCATTTGTATGAGGAGCGGCACATGATCATGCATACTCCCGGGTGGTGCCCT	216
182	CGCTTCGGGAGGCAACTCTCCCTTGTGTTGGTGGCGCTCACTCCACGCTCGCGGCCA	241
217	CGCTTCGGGAGCAACAAGCTCCCGTTGCTGGTAGCGTCACTCCACGCTCGCGGCCA	276
242	GGAAACGCAGCGTCCCAACAACAATACGACGCCACGTCGATTTGCTGTTGGGGCTG	301
277	GGAAATGCCAGCGTCCCCACTACGACAATAACGACGCCACGTCGACTTGTCTGTTGGGGCG	336
302	CTGCTTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTTCCTGTTTCC	361
337	CTGCTTTTCTGCTTCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTTCCTGCTCCG	396
362	AGCTGTTTCACTTTCTCACTCGCGCGGCATCAAAACAGTACGAGGACTGCAATCTCAATCT	421
397	AGCTGTTTCACTTTCTGCTCTCGCGCGCATGACAGTGCAGGACTGCAATCTCAATCT	456
422	ATCCCGGGCCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGA	466
457	ATCCCGGGCCATTTATCAGGTCAACCGCATGGCTTGGGATATGATGA	501

RESULT 14

US-08-612-973-3
; Sequence 3, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.

```

;
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..639
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: 1..636
;
US-08-612-973-3

```

Query Match	88.0%	Score 424.8	DB 3	Length 642
Best Local Similarity	93.3%	Pred. No. 2.8e-110		
Matches 444; Conservative	0	Mismatches 32	Indels 0	

Qy	1	ATGCCCGGTGCTCTTTTCTCTATCTTCTCTTTGGCCCTGCTGTCCTGTCTGACCAATACCA	60
Db	1	ATGCCCGGTGCTCTTTTCTCTATCTTCTCTTTGGCTTTACTGTCTGTCTGACCAATCCA	60
Qy	61	GCTTCCGCTTATGAAGTGCGCCAAAGTGTCCGGGGTGTAACCATGTCCAGAACGACTGCTCC	120
Db	61	GCTTCCGCTTATGAGTGGCGCAAAGTGTCCGGGATGTACCATGTCCAGAACGACTGCTCC	120
Qy	121	AACTCAAGCATPATGTATGAGGCAGCGGACATGATCATGCAACCCCGGGTGCGTGCC	180
Db	121	AACTCAAGCATPTGTATGAGGCAGCGGACATGATCATGCAACCCCGGGTGCGTGCC	180
Qy	181	TGCGTTCCGGGAGGCGCACTCTCCGTTCGTGGGTGGCGCTCACTCCACGCTCCGGCC	240
Db	181	TGCGTTCCGGGAGAACACTCTTCCCGCTGCTGGGTAGCGCTCAACCCACGCTCGGAGCT	240
Qy	241	AGGAACGCCACGCGTCCCCACAACGACAATACGACGCCACGTCGATTTGCTCTTGGGGCT	300
Db	241	AGGAACGCCACGCGTCCCCACCACGACAATACGACGCCACGTCGATTTGCTCTTGGGGCG	300
Qy	301	GCTGCTTCTGTTCCGCTATGTAGTGGGGGATCTTCGGGATCTGTTTTCTTGTTTCC	360
Db	301	GCTGCTCTCTGTTCCGCTATGTAAGTGGGGGATCTTCGGGATCTGTCTTCCTCGTCTCC	360
Qy	361	CAGCTGTTCACTTCTCACTTCGCGGGCATCAACAGTACAGGACTGCAACTGCTCAATC	420
Db	361	CAGCTGTTCAACATCTCGCCTCGCGGGCATGAGACGGTGCAGGACTGCNAATGCTCAATC	420
Qy	421	TATCCCGGCCATGTATCAGGTCAACGCATGCGCTTGGGATATGATGATGAATCGTGC	476
Db	421	TATCCCGGCCACATAACAGGTCAACGTTGGCTTGGGATATGATGATGAATCGTGC	476

RESULT 15

RES001 13
 US-08-927-597-3
 : Sequence 3, Application US/08927597
 : Patent No. 6245503
 : GENERAL INFORMATION:
 : APPLICANT: MAERTENS, GERT
 : APPLICANT: BOSMAN, FONS
 : APPLICANT: DE MARTYNOFF, GUY
 : APPLICANT: BUYSSE, MARIE-ANGE
 : TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
 : TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
 : NUMBER OF SEQUENCES: 111
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: NIXON & VANDERHVE P. C.
 : STREET: 1100 NORTH GLEBE ROAD
 : CITY: ARLINGTON
 : STATE: VIRGINIA
 : COUNTRY: U.S.A.
 : ZIP: 22201-4714
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/927-597

Mon Dec 22 13:28:54 2003

QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	
1	ATGCCCGGTGCTCTTTCTCTATCTTCTCTGTGGCCCTGCTGTCTGCTGTGACCATCA	60	1	ATGCCCGGTGCTCTTTCTCTATCTTCTCTGTGGCCCTGCTGTCTGCTGTGACCATCA	60	61	GCTTCCGCTTATGAAGTCGCGCAACGTGTCCGGGTGTACCATGTCAAGACGACTGCTCC	120	61	GCTTCCGCTTATGAAGTCGCGCAACGTGTCCGGGTGTACCATGTCAAGACGACTGCTCC	120	121	AACCTAAGCATATGTATGAGCAGCGACATGATCATGCACACCCCGGTGGCTGCC	180	121	AACCTAAGCATATGTATGAGCAGCGACATGATCATGCACACCCCGGTGGCTGCC	180	
181	TGCGTTTCGGAGGGCAACTCCCTCCGCTTGTGGGTGGCGCTCACTCCACGCTCGCGCC	240	181	TGCGTTTCGGAGGAAACAACTCTTCCGCTGCTGGTAGGCTCACCCCAACGCTCGCAGCT	240	241	AGGAACGCCAGCGTCCCAACAAAGACATACAGCCGACGATTCGTCTGTTGGGCT	300	241	AGGAACGCCAGCGTCCCAACCAATACAGCCGACGATTCGTCTGTTGGGCT	300	301	GCTGCTTTCTGTTCGCTATGTACGTGGGGATCTCTGCGGATCTGTTTCCTGTTTC	360	301	GCTGCTCTCTGTTCCGCTATGTACGTGGGGATCTCTGCGGATCTGTTTCCTGTTTC	360	
361	CAGCTGTTCACTTCTCACTCCCGCGCATCAAAAGTACAGGACTGCAACTGCTCAATC	420	361	CAGCTGTTCAACCATCTCGCTCCCGCATGAGACGCTGCAGGACTGCAATTCCTCAATC	420	421	TATCCCGCCCATGTATCAGTCAACGATCCGCTTGGGATATGATGAACTGGTC	476	421	TATCCCGCCCATTAACAGGTACCGTATGCTTGGGATATGATGAACTGGTC	476							

Search completed: December 20, 2003, 07:03:00
Job time : 35.8338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 1817.41 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-11
Perfect score: 480
Sequence: 1 ATGTCGGTGTCTTCTC.....TGATGATGAACGGTAATAG 480

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
- 1: gb_ba.*
 - 2: gb_hgt.*
 - 3: gb_in.*
 - 4: gb_om.*
 - 5: gb_ov.*
 - 6: gb_pat.*
 - 7: gb_ph.*
 - 8: gb_pl.*
 - 9: gb_pr.*
 - 10: gb_ro.*
 - 11: gb_sts.*
 - 12: gb_sy.*
 - 13: gb_un.*
 - 14: gb_vi.*
 - 15: em_ba.*
 - 16: em_fun.*
 - 17: em_hum.*
 - 18: em_in.*
 - 19: em_mu.*
 - 20: em_om.*
 - 21: em_or.*
 - 22: em_ov.*
 - 23: em_pat.*
 - 24: em_ph.*
 - 25: em_pl.*
 - 26: em_ro.*
 - 27: em_sts.*
 - 28: em_un.*
 - 29: em_vi.*
 - 30: em_htg_hum.*
 - 31: em_htg_inv.*
 - 32: em_htg_other.*
 - 33: em_htg_mus.*
 - 34: em_htg_pln.*
 - 35: em_htg_rod.*
 - 36: em_htg_man.*
 - 37: em_htg_vrt.*
 - 38: em_sy.*
 - 39: em_htgo_hum.*
 - 40: em_htgo_mus.*
 - 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	480	100.0	480	6	A48673	A48673 Sequence 11
2	480	100.0	480	6	AR157328	AR157328 Sequence
3	480	100.0	480	6	AX452760	AX452760 Sequence
4	480	100.0	480	6	AX685012	AX685012 Sequence
5	474.2	98.8	483	6	A48671	A48671 Sequence 9
6	474.2	98.8	483	6	AR157327	AR157327 Sequence
7	474.2	98.8	483	6	AX452758	AX452758 Sequence
8	474.2	98.8	483	6	AX685010	AX685010 Sequence
9	436	90.8	9379	14	AF207766	AF207766 Hepatitis C
10	432.8	90.2	1880	14	HPCSTRJ4	D00832 Hepatitis C
11	432.8	90.2	2540	6	E04260	E04260 CDNA encodi
12	432.8	90.2	2540	6	E04805	E04805 CDNA to 5'-
13	432.8	90.2	2540	6	E07391	E07391 cDNA encodi
14	432.8	90.2	9448	14	HPCJ483	D13558 Hepatitis C
15	431.2	89.8	1539	6	AR027786	AR027786 Sequence
16	431.2	89.8	1863	6	AR027783	AR027783 Sequence
17	430.8	89.8	9580	14	AF054250	AF054250 Hepatitis
18	429.6	89.5	8780	14	AF054257	AF054257 Hepatitis
19	429.6	89.5	8780	14	AF054258	AF054258 Hepatitis
20	428	89.2	8781	14	AF054255	AF054255 Hepatitis
21	428	89.2	8781	14	AF054254	AF054254 Hepatitis
22	428	89.2	9379	14	AF165052	AF165052 Hepatitis
23	428	89.2	9460	14	HPCJ491	D10750 Hepatitis C
24	428	89.2	9595	6	AR119832	AR119832 Sequence
25	428	89.2	9595	14	AF054247	AF054247 Hepatitis
26	428	89.2	9596	14	AF054249	AF054249 Hepatitis
27	428	89.2	9599	6	AR119833	AR119833 Sequence
28	426.4	88.8	577	6	E04085	E04085 gDNA encodi
29	426.4	88.8	618	14	HPCJ491	D28929 Hepatitis C
30	426.4	88.8	8779	14	AF054251	AF054251 Hepatitis
31	426.4	88.8	8780	14	AF054252	AF054252 Hepatitis
32	426.4	88.8	9410	14	HPCJ182	D50481 Hepatitis C
33	426.4	88.8	9595	14	AF054248	AF054248 Hepatitis
34	426.2	88.8	633	6	A48669	A48669 Sequence 7
35	426.2	88.8	633	6	AR157326	AR157326 Sequence
36	426.2	88.8	633	6	AX452756	AX452756 Sequence
37	426.2	88.8	633	6	AX685008	AX685008 Sequence
38	426	88.8	9377	14	AF207756	AF207756 Hepatitis
39	426	88.8	9435	14	AB049093	AB049093 Hepatitis
40	424.8	88.5	3296	14	AB008446	AB008446 Hepatitis
41	424.8	88.5	8780	14	AF054259	AF054259 Hepatitis
42	424.8	88.5	8781	14	AF054256	AF054256 Hepatitis
43	424.8	88.5	9361	14	AF483269	AF483269 Hepatitis
44	424.8	88.5	9369	14	AF165054	AF165054 Hepatitis
45	424.8	88.5	9379	14	AF165051	AF165051 Hepatitis

ALIGNMENTS

RESULT 1
A48673
LOCUS A48673 Sequence 11 from Patent WO9604385. 480 bp DNA linear PAT 07-MAR-1997
DEFINITION A48673
ACCESSION A48673
VERSION A48673.1 GI:2302386
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 480)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 11 15-FEB-1996;

BASE COUNT	85 a	149 c	123 g	123 t
ORIGIN				
Query Match	100.0%; Score 480; DB 6; Length 480;			
Best Local Similarity	100.0%; Pred. No. 5e-113;			
Matches 480; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGTCCGGTTCCTTTCTCTATCTTCTCCTCTGTTGGCCCTGCTGTCTCTGACCATACCA	60	
DB	1	ATGTCCGGTTCCTTTCTCTATCTTCTCCTCTGTTGGCCCTGCTGTCTCTGACCATACCA	60	
QY	61	GCTTCCGCTTATGAAGTGCGCAACGTCGCGGGGTACCATGTCACGACGACTGCCTCC	120	
DB	61	GCTTCCGCTTATGAAGTGCGCAACGTCGCGGGGTACCATGTCACGACGACTGCCTCC	120	
QY	121	AATCTAAGCATAGTGTATGAGCGACGCCGACATGATCATGCACACCCCGGGTGGTGGCC	180	
DB	121	AATCTAAGCATAGTGTATGAGCGACGCCGACATGATCATGCACACCCCGGGTGGTGGCC	180	
QY	181	TGGCTTCGGAGGGCAACTCCTCCGTTGCTGGGTGGCGTCACTCCCACGCTCGCGGCC	240	
DB	181	TGGCTTCGGAGGGCAACTCCTCCGTTGCTGGGTGGCGTCACTCCCACGCTCGCGGCC	240	
QY	241	AGNAGCCGACGCTCCCCACAACAGCAATACGACGCCACGTCGATTTGCTGTTGGGGCT	300	
DB	241	AGNAGCCGACGCTCCCCACAACAGCAATACGACGCCACGTCGATTTGCTGTTGGGGCT	300	
QY	301	GCTGCTTTCGTTCCCGTATGTAGTCGTCGGGGATCTCTGCGGATCTGTTTCCTGTTTCC	360	
DB	301	GCTGCTTTCGTTCCCGTATGTAGTCGTCGGGGATCTCTGCGGATCTGTTTCCTGTTTCC	360	
QY	361	CAGCTGTTCACTTCTCACCTCGCGGCATCAAACAGTACAGGACTGCAACTGCTCAATC	420	
DB	361	CAGCTGTTCACTTCTCACCTCGCGGCATCAAACAGTACAGGACTGCAACTGCTCAATC	420	
QY	421	TATCCCCGGCATGTATCAGTCAACGCGATGGCTTGGGATATGATGATGAAGTGTAAATAG	480	
DB	421	TATCCCCGGCATGTATCAGTCAACGCGATGGCTTGGGATATGATGATGAAGTGTAAATAG	480	
RESULT 4				
AX685012	480 bp DNA linear PAT 29-MAR-2003			
LOCUS	Sequence 11 from Patent WO02055548.			
DEFINITION	AX685012			
ACCESSION	AX685012.1 GI:29371417			
VERSION	Hepatitis C virus			
KEYWORDS	Hepatitis C virus			
SOURCE	Hepatitis C virus			
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
REFERENCE	1			
AUTHORS	Maertens,G., Bosman,F. and Buysse,M.A.			
TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use			
JOURNAL	Patent: WO 02055548-A 11 18-JUL-2002;			
FEATURES	INNOGENETICS N.V. (BE)			
source	Location/Qualifiers			
	1..480			
	/organism="Hepatitis C virus"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:11103"			
	1..477			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAD86523.1"			
	/db_xref="GI:29371418"			
	/translation="MSGCSFSPILLALLSLCTIPASAVEVRNVGVVHTNDNCSSNI			
	VYEADIMMTPGCVPSQRGNSSRCVALPTLTAARNASVPPTTIIRHVDDLLVGAAAL			
	FCSAWYVDGLCGSVFLVSQJFTFSRRRHTQVDCNCNSIYPGHVSGHRMAWMNNW"			
	1..474			
	/product="unnamed"			
CDS				
mat_peptide				

BASE COUNT	85 a	152 c	123 g	123 t
ORIGIN				
Query Match	98.8%;	Score 474.2;	DB 6;	Length 483;
Best Local Similarity	99.4%;	Pred. No. 1.6e-111;		
Matches 476;	Conservative	0; Mismatches	3; Indels	0; Gaps
0;				
QY	1	ATGTCGGGTGCTCTTTCTCTATCTCTCTCTTTGGCCCTGCTGTCTGTCTGTGACCATACCA	60	
Db	1	ATGCCCGGTGTGCTCTTTCTCTATCTCTCTCTTTGGCCCTGCTGTCTGTCTGTGACCATACCA	60	
QY	61	GCTTCCGCTTATGAAGTGCACAACTGTCGCGGGTGTAACATGTACAGAACGACTGCTCC	120	
Db	61	GCTTCCGCTTATGAAGTGCACAACTGTCGCGGGTGTAACATGTACAGAACGACTGCTCC	120	
QY	121	AACCTAAGCATAGTGTATGAGGACGCGGACATGATCATGCACACCCCGGGTGGGTGCCC	180	
Db	121	AACCTAAGCATAGTGTATGAGGACGCGGACATGATCATGCACACCCCGGGTGGGTGCCC	180	
QY	181	TGGTTTCGGAGGCGCAACTCTCCCGTTGCTGGGTGGCGTCACTCCACGCTCGCGGCC	240	
Db	181	TGGTTTCGGAGGCGCAACTCTCCCGTTGCTGGGTGGCGTCACTCCACGCTCGCGGCC	240	
QY	241	AGAAACCCAGCGTCCCAACGACCAATACGACGCCAGTCGATTTGCTCGTTGGGGCT	300	
Db	241	AGAAACCCAGCGTCCCAACGACCAATACGACGCCAGTCGATTTGCTCGTTGGGGCT	300	
QY	301	GCTGCTTTCGTTCCCGCTATGTACGTGGGGATCTCTCGCGATCTGTTTTCCTTGTTC	360	
Db	301	GCTGCTTTCGTTCCCGCTATGTACGTGGGGATCTCTCGCGATCTGTTTTCCTTGTTC	360	
QY	361	CAGCTGTTCACCTTCTCACTCCGCGCATCAAAAGTACAGGACTGCAACTGCTCAATC	420	
Db	361	CAGCTGTTCACCTTCTCACTCCGCGCATCAAAAGTACAGGACTGCAACTGCTCAATC	420	
QY	421	TATCCCGGCATGATCAGTTCACCGCATGGCTTGGGATGATGATGAATCGTGAATA	479	
Db	421	TATCCCGGCATGATCAGTTCACCGCATGGCTTGGGATGATGATGAATCGTGAATA	479	
RESULT 6				
AR157327				
LOCUS	AR157327	483 bp	DNA	linear
DEFINITION	Sequence 9 from patent US 6245503.			PAT 17-OCT-2001
ACCESSION	AR157327.1			
VERSION	GI:16218260			
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 483)			
AUTHORS	Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.			
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use			
JOURNAL	Patent: US 6245503-A 9 12-JUN-2001;			
FEATURES	Location/Qualifiers			
source	1..483			
	/organism="unknown"			
BASE COUNT	85 a	152 c	123 g	123 t
ORIGIN				
Query Match	98.8%;	Score 474.2;	DB 6;	Length 483;
Best Local Similarity	99.4%;	Pred. No. 1.6e-111;		
Matches 476;	Conservative	0; Mismatches	3; Indels	0; Gaps
0;				
QY	1	ATGTCGGGTGCTCTTTCTCTATCTCTCTCTTTGGCCCTGCTGTCTGTGACCATACCA	60	
Db	1	ATGCCCGGTGTGCTCTTTCTCTATCTCTCTCTTTGGCCCTGCTGTCTGTGACCATACCA	60	
QY	61	GCTTCCGCTTATGAAGTGCACAACTGTCGCGGGTGTAACATGTACAGAACGACTGCTCC	120	
Db	61	GCTTCCGCTTATGAAGTGCACAACTGTCGCGGGTGTAACATGTACAGAACGACTGCTCC	120	

[illegible]

GVNATCNLPGCSFSLFLLALISCLTIPASAVEVRNVSGIYHVNDSCSSIVYEAD
MIMHTPCVPCVEDNNSRCWALPTFLAARNASVPTTIRRHVDLLVGAAPCSAMY
VDGLGSLFVLSQLETFSPRHETVQDCNSIYPGHLGSRHAWDMNMSPTTALV
SOLLRIPOAVDMVAGAHGVLGLAYYGMVGNKAVLIVALLFAGVGETYTGAA
SHTTSLASLFPSPGASORIQLVNTGNSWHNRTALNCNLSLHTGFLAALPYTHRFNS
GCPERMASCRPIDWFAQGWGPITYTEPDSPDQRPYCMHYAPRPGIVPASPQVCGPYC
FTPSP"

BASE COUNT 333 a 593 c 555 g 399 t

Query Match 90.2%; Score 432.8; DB 14; Length 1880;

Best Local Similarity 94.3%; Pred. No. 7.4e-101;

Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY	2	TGTCGGTGTCTTTCTCTATCTTCTCTTGTGCGCCCTGCTGCTGTCTGACCATACCAG	61
DB	847	TGCGCGGTGTCTTTCTCTATCTTCTCTTGTGCTTGTGCTGCTGCTGTGACCATCCAG	906
QY	62	CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTGTACCATGTACGACGAGCTGTCTCA	121
DB	907	CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTATACCATGTACGACGAGCTGTCTCA	966
QY	122	ACTCAAGCATAGTGTATGAGCGAGCGGACATGATCATGTACACACCCCGGGTGGTGCCT	181
DB	967	ACTCAAGCATAGTGTATGAGCGAGCGGACATGATCATGTACACACCCCGGGTGGTGCCT	1026
QY	182	GGTTTCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA	241
DB	1027	GGTTTCGGAGGAGCAACAGCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA	1086
QY	242	GGAGCCGAGCGTCCCGACACACATAGCAGCCACGCTGATTTGCTGTTGGGGCTG	301
DB	1087	GGAGCCGAGCGTCCCGACACATAGCAGCCACGCTGATTTGCTGTTGGGGCTG	1146
QY	302	CTGCTTTCTGCTCGGTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC	361
DB	1147	CTGCTTTCTGCTCGGTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC	1206
QY	362	AGCTGTTCACCTTCTCACCTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT	421
DB	1207	AGCTGTTCACCTTCTCACCTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT	1266
QY	422	ATCCCGGCGCATGTATCAGTCAAGTCCCGCATCGCTTGGGATATGATGAACTGTGTAA	477
DB	1267	ATCCCGGCGCATTTATCAGTCAAGTCCCGCATCGCTTGGGATATGATGAACTGTGTAA	1322

RESULT 11

E04260
LOCUS
CDNA encoding a part of type non-A non-B hepatitis virus.
PAT 29-SEP-1997

E04260

VERSION
E04260.1 GI:2172463

KEYWORDS
JP 1993023200-A/2.

SOURCE
unidentified

ORGANISM
unclassified.

REFERENCE
1 (bases 1 to 2540)

AUTHORS
Okamoto, H. and Nakamura, T.

TITLE
HIGHLY SENSITIVE DETECTION METHOD OF NON-A NON-B TYPE HEPATITIS

JOURNAL
VIRUS USING OLIGONUCLEOTIDE PRIMER AND OLIGONUCLEOTIDE PRIMER

COMMENT
Patent: JP 1993023200-A 2 02-FEB-1993;

NAKAMURA TETSUO

PN JP 1993023200-A/2

PD 02-FEB-1993

PF 26-FEB-1991 JP 1991191376

PR 12-JUN-1990 JP 90P 153402

PI OKAMOTO HIROAKI, NAKAMURA TETSUO

PC C12Q1/68.C12N15/51.C12Q1/70;

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

*source: strain=HC-J4;
Key Location/Qualifiers

misc_feature 1..2540

FT notes='a part of type non-A non-B hepatitis virus';

Location/Qualifiers

1..2540

/organism='unidentified'

/mol_type='genomic RNA'

/db_xref='taxon:32644'

BASE COUNT 470 a 776 c 741 g 553 t

ORIGIN

Query Match	90.2%;	Score 432.8;	DB 6;	Length 2540;
Best Local Similarity	94.3%;	Pred. No. 7.4e-101;		
Matches 449;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
QY	2	TGTCGGTGTCTTTCTCTATCTTCTCTTGTGCGCCCTGCTGCTGTCTGACCATACCAG	61	
DB	847	TGCGCGGTGTCTTTCTCTATCTTCTCTTGTGCTTGTGCTGCTGTGACCATCCAG	906	
QY	62	CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTGTACCATGTACGACGAGCTGTCTCA	121	
DB	907	CTTCCGCTTATGAAGTGGCGCAACGTCTCGGGGTATACCATGTACGACGAGCTGTCTCA	966	
QY	122	ACTCAAGCATAGTGTATGAGCGAGCGGACATGATCATGTACACACCCCGGGTGGTGCCT	181	
DB	967	ACTCAAGCATAGTGTATGAGCGAGCGGACATGATCATGTACACACCCCGGGTGGTGCCT	1026	
QY	182	GGTTTCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA	241	
DB	1027	GGTTTCGGAGGAGCAACAGCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA	1086	
QY	242	GGAGCCGAGCGTCCCGACACAAATAGCAGCCACGCTGATTTGCTGTTGGGGCTG	301	
DB	1087	GGAGCCGAGCGTCCCGACCAATAGCAGCCACGCTGATTTGCTGTTGGGGCTG	1146	
QY	302	CTGCTTTCTGCTCGGTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC	361	
DB	1147	CTGCTTTCTGCTCGGTATGATGAGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC	1206	
QY	362	AGCTGTTCACCTTCTCACCTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT	421	
DB	1207	AGCTGTTCACCTTCTCACCTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT	1266	
QY	422	ATCCCGGCGCATGTATCAGTCAAGTCCCGCATCGCTTGGGATATGATGAACTGTGTAA	477	
DB	1267	ATCCCGGCGCATTTATCAGTCAAGTCCCGCATCGCTTGGGATATGATGAACTGTGTAA	1322	

RESULT 12

E04805

LOCUS

CDNA to 5'-terminal region of grNA of Hepatitis nonA nonB virus.

DEFINITION

E04805

ACCESSION

VERSION

E04805.1 GI:2173001

KEYWORDS

JP 1993091884-A/2.

SOURCE

unidentified

ORGANISM

unclassified.

REFERENCE

1 (bases 1 to 2540)

AUTHORS

Okamoto, H. and Nakamura, T.

TITLE

DETECTION SYSTEM FOR NON-A NON-B HEPATITIS VIRUS RELATING ANTIGEN

JOURNAL

AND ANTIBODY, POLYNUCLEOTIDE AND POLYPEPTIDE

COMMENT

Patent: JP 1993091884-A 2 16-APR-1993;

NAKAMURA TETSUO

OS Hepatitis nonA nonB virus

PN JP 1993091884-A/2

PD 16-APR-1993

PF 10-APR-1991 JP 1991196175

PR 12-JUN-1990 JP 90P 153401,

OKAMOTO HIROAKI, NAKAMURA TETSUO

304405 PI

OS	Hepatitis non-A non-B virus
PN	JP 1994125777-A/4
PD	10-MAY-1994
PF	20-JUN-1991 JP 1991247120
PI	OKAMOTO HIROAKI, NAKAMURA TETSUO
PC	C12N15/51.C12N1/21.C12P21/02.C12Q1/70.G01N33/53.G01N33/569, PC G01N33/576;
CC	strandedness: Single;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No; Location/Qualifiers
FH	Key
FT	source
FT	1..2540
FT	/organism='Hepatitis non-A non-B virus' FT
FT	/strains='HC-J4'
FT	5'UTR 1..341
FT	mat_peptide 342..1490
FT	/note='non-structural protein of chimpanzee
FT	hepatitis non-A
FT	non-B virus'
FT	mat_peptide 1491..2540
FT	/note='Structural protein of chimpanzee FT
FT	hepatitis non-A
FT	non-B virus'.
FEATURES	
source	Location/Qualifiers
	1..2540
	/organism='unidentified'
	/mol_type='genomic RNA'
	/db_xref='taxon:32644'
BASE COUNT	470 a 776 c 741 g 553 t
ORIGIN	
Query Match	90.2%; Score 432.8; DB 6; Length 2540;
Best Local Similarity	94.3%; Pred. No. 7.4e-101;
Matches 449; Conservative	0; Mismatches 27; Indels 0; Gaps 0;
Qy	2 TGTCCGGTGTCTTCTCTATCTTCCTCTTGGCCCTGTCTCTCTCTGACCATACCAG 61
Db	847 TGCCTGGTGTCTTCTCTCTATCTCTCTTGGCTTGGCTCTCTGTGTTGACCATCCAG 906
Qy	62 CTTCCGCTTATGAAGTGGCCAAAGTGTCCGGGTGTACCATGTCAAGACGACTGTCCA 121
Db	907 CTTCCGCTTATGAAGTGGCCAAAGTGTCCGGGTATATACCATGTCAAGACGACTGTCCA 966
Qy	122 ACTCAAGCATAGTGTATGAGGAGGACGATGATCATGCACACCCCGGTGCGTCCCT 181
Db	967 ACTCAAGCATAGTGTATGAGGAGGACGATGATCATGCATCTCTCCGGTGTGCGCT 102
Qy	182 GCGTTCCGGAGGGCAACTCCTCCCGTGTCTGGGTGGGCTCATTCCACGCTCGCGCCA 241
Db	1027 GCGTTCCGGAGGACAAAGCTCCCGTGTCTGGGTAGCGCTCACTCCACGCTCGCGCCA 108
Qy	242 GGAACGCCAGCTGCCCAACAGCAATACGACGCGCAGTGCATTTGCTGTTGGGGCTG 301
Db	1087 GGAATGCCAGCGTCCCACTACGACAAATACGACGCGCAGTGCATTTGCTGTTGGGGCTG 114
Qy	302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTGTCTTCTGTTTCCC 361
Db	1147 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTGTCTTCTGTTTCCC 120
Qy	362 AGCTGTTTCACTTCTCACTCCGGGCATCAACAGTACAGACTGCAACTGCTCAATCT 421
Db	1207 AGCTGTTTCACTTCTCACTCCGGGCATCAACAGTACAGACTGCAACTGCTCAATCT 126
Qy	422 ATCCCGGCCATGTATCAGGTACCGCATGGCTTGGGATATGATGAATCGTAA 477
Db	1267 ATCCCGGCCATTTATCAGGTACCGCATGGCTTGGGATATGATGAATCGTAA 1322
RESULT 14	
HPCJ483	
LOCUS	HPCJ483 9448 bp RNA linear VPL 01-FEB-2001

DEFINITION	Hepatitis C virus genome, complete sequence.
ACCESSION	D13558 D01217
VERSION	D13558.1 Gi:221604
KEYWORDS	C protein; E protein; E2 protein; NS1 protein; NS2 protein; NS3 protein; NS4 protein; NS5 protein.
SOURCE	Hepatitis C virus
ORGANISM	Hepatitis C virus
REFERENCE	1 (bases 1 to 9448) Okamoto, H., Kojima, M., Okada, S., Yoshizawa, H., Iizuka, H., Tanaka, T., Muchmore, E.E., Peterson, D.A., Ito, Y. and Mishiro, S. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability Virology 190 (2), 894-899 (1992)
AUTHORS	Okamoto, H.
TITLE	Direct Submission
JOURNAL	Submitted (17-OCT-1991) Hiroaki Okamoto, Jichi Medical School, Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi 329-04, Japan (E-mail: hokamoto@jichi.ac.jp, Tel: 0285-44-2111 (ex 3334), Fax: 0285-44-1557)
MEDLINE	Submitted (17-Oct-1991) to DDBJ by:
PUBMED	Hiroaki Okamoto Immunology Division Jichi Medical School Kawachi-gun Tochigi 329-04 Japan Phone: 0285-44-2111 x3334 Fax: 0285-44-1557.
REFERENCE	1. .9448 Location/Qualifiers /organism="Hepatitis C virus" /mol_type="genomic RNA" /strain="HC-J4" /isolates="HC-J4/83" /db_xref="taxon:11103" 342_..9374 /codon_start=1 /product="polyprotein precursor" /protein_id="BAA02756.1" /db_xref="GI:221605"
AUTHORS	/translation="MSTNPKPQKTKRNTNRNRPQDVKFPFGGQIVGGVLLPRGPRL GVRATEKTSERQPRGWRQPIPKARPEGRWAQPCYPMPLVNGNEGLHWAGLLSPRG SRPSGPTDPRKSRNIGKVIDLTCTGFADLMGIIPLVGAFLGGAARAAHGVRLLED GVNYATNPUGPCSFSIFLLALLSLTIPASAVEVRNVSGIYHVTNDCSNISYVEAAD MIMHTPGCVPCVRNEDSSRCWALPTLAARNASVPTTIRRHVDLLVGAAAFCSAMY VGDLCGSLVLSOLFSPRRHETVQDCNCSYPGHLSGRHMAWMMWMSPTTALV SOLLRTPOAVDMVAGAHVGLAGLAYYSMVGNAKVLIVALLFAGVDCGETVSGAA SHTTSLASLSPGASQRIQLVNTNGSWHINRTALNCNLSLHGTFLAALPYTHRFNS GCPERNASCRPIIDWFAQGGPIITYEPDSDQRPICWYAPRCPIVPSAQVCGPVYC FPTSPVVTGTRDSGTYSWGENETDVMNNRPNQGNWFGCTWNNSTYTKTCGG PCNTGVGNHTLTCTCFRKHPEATYTKCGSPWLTPLRCLVDYVYRLWHYPTNF SIFKVMYVGGVHRLNAACTNRGRCNLEDRSELSPLLSTTEWOLLPCAFTTL PALSTGLIHLHONIVDQVLYGVGSFVFAIKWEYIILLFLLADARVACIWMMLL IAQAEALNVLVNAASVAGAHGILSLVFFCAWYIKGR LAPGAAYAFYGVWPLL LLALLPRAYAMDREMAASCGGAVLGLVLTLSPIYKVLTKLILWLOFTITRAEAB MQVWPLNVRGDRDAIILLTCAVHPELFDITKLALLILGLMLVQLQITRVPYVR AAGLIRACMLKLVACGHVQWAFMKLGALTGYVYNNHLTPLRDMAHAGRLDLAVE PVVFSDMETKVTIWTGADTAACGDIILGLPVSARRGKEIFLGPADSLEGOWRLLAIT AYSQOVRGLVLCITSLTGRDKNOVEGVVSTATQSLATCINGVCTVTVHAGSK TLAGPKGPIQMTYNTVDLILVGNQAPPGARSMTPTCGSDSLVLRHADVIVRRRG DSRGLSPRPVYLKSGSGGLLPCSPGVGVGVVFRAACTRGVAKADFIPIVSMETT MRSPTDKNSSPPVQPTQVAHLHAPTSGSKTKVPAAYAGYKVLVNLNPSVAATL GFGAMPKAGHDIPNRTGVTITITGGPIITYSTYKFLADGGCGGAYDIIIDECIS TOSTTILGTLGTDOATAGARLVLATATPPGSVTVPHNTEIEGLSNNGEIPFYGK AIPLEAKGSRHILFCHSKKCDLAKLTGLGNAYVYRGLDVSVIPIQGDVAVA TDALMTGFGDFNSVDNCNTCVQTQVDFSLDPTFTTETTTVPDASVRSQRRTGVR RSGIYRFTVPGRPSGMFSSVLCEDYDAGCAWYELTPAETSVLRLEYLNTPLGPVQC DHLFEWESVFTGLSHIDAHFLSQTKAGDNFPYLVAYQATVCARAQAPPSPWDQWKC
FEATURES	source
CDS	mat_peptide 342..914 /product="C protein" mat_peptide 915..1490 /product="E protein" mat_peptide 1491..2528 /product="NS1/E2 protein" mat_peptide 2529..3359 /product="NS2 protein" mat_peptide 3360..5186 /product="NS3 protein" mat_peptide 5187..6380 /product="NS4 protein" mat_peptide 6381..9371 /product="NS5 protein" stem_loop 6005..6104 /note="possible stem-loop structure in NS4 sequence" misc_feature 9413..9448 /note="T-stretch of 36 bp" BASE COUNT 1899 a 2828 c 2676 g 2045 t ORIGIN
	Query Match 90.2%; Score 432.8; DB 14; Length 9448; Best Local Similarity 94.3%; Pred. No. 7.6e-101; Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0; QY 2 TGTCCGGTTGCTCTTTCTCTATCTTCTTGGCCCTGCTGCTCTGTGACCATACCAG 61 DB 847 TGCCCGGTTGCTCTTTCTCTATCTTCTTGGCTTGTGCTGTTGACCATCCAG 906 QY 62 CTTTCGGCTTATGAAGTGCACACGCTGTCGGGGTGATCCATGTACGAAACGACTGTCTCA 121 DB 907 CTTTCGGCTTATGAAGTGCACACGCTGTCGGGGATATACCATGTGACGAAACGACTGTCTCA 966 QY 122 ACTCAACGATAGTGTATGAGGACGCGGACATGATCATGACACACCCCGGGTGGTGGCCT 181 DB 967 ACTCAACGATAGTGTATGAGGACGCGGACATGATCATGATCATCTCCCGGGTGGTGGCCT 1026 QY 182 GCCTTCGGAGGGCAACTCTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCCA 241 DB 1027 GCCTTCGGAGGCAACACGCTCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGGCCA 1086 QY 242 GGAACGCCAGCGTCCCAACACGACAAATACGACGCCACGCTCGATTTGCTGTTGGGGCTG 301 DB 1087 GGAATGCCAGCGTCCCAACACGACAAATACGACGCCACGCTCGATTTGCTGTTGGGGCTG 1146 QY 302 CTGCTTTCTGTTCCGCTATGACGCGGGGATCTCTGCGGATCTGTTTCTGTTTCTGTTTCCC 361 DB 1147 CTGCTTTCTGCTCCGCTATGACGCGGGGATCTCTGCGGATCTGTTTCTGTTTCTGTTTCCC 1206

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 132.7 Seconds

(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-11

Perfect score: 480

Sequence: 1 ATGTCGGTGTGCTCTTCTC.....TGATGATGAACCTGTAATAG 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	17	AAT12708
2	480	100.0	480	24	AA148917
3	474.2	98.8	483	17	AAT12707
4	474.2	98.8	483	24	AA148916
5	432.8	90.2	1880	13	AA024467
6	432.8	90.2	2187	19	ABA03491
7	432.8	90.2	2540	14	AA043889
8	432.8	90.2	2540	15	AA063753

9	431.2	89.8	1863	12	AAQ15363	Fragment of NANB H
10	429.6	89.5	2540	13	AAQ29628	Hepatitis C virus
11	428	89.2	9595	20	AA24843	Infectious hepatitis
12	428	89.2	9595	22	AA086939	Nucleotide sequenc
13	428	89.2	9595	22	AA086939	Infectious hepatitis
14	428	89.2	9599	20	AA24833	Infectious hepatitis
15	426.4	88.8	577	14	AAQ35081	HCV envelope regio
16	426.4	88.8	2187	19	ABA03492	Cuticle protein 1
17	426.2	88.8	633	17	AAT12706	HCV E1 construct H
18	426.2	88.8	633	24	AA148915	Hepatitis C virus
19	424.6	88.5	567	13	AAQ27160	NANB hepatitis vir
20	423.2	88.2	577	14	AAQ35085	HCV envelope regio
21	423.2	88.2	580	12	AAQ11076	Fragment of hepati
22	423.2	88.2	580	20	AAZ07647	HCV J1 E domain co
23	423.2	88.2	1562	19	AAV60672	Fragment #5 isolat
24	423.2	88.2	1953	25	AA155222	Plasmid pDK2 DNA
25	423.2	88.2	2829	19	AAV60673	Fragment #6 isolat
26	422.8	88.1	580	16	AAQ79750	Hepatitis C virus
27	422.8	88.1	580	20	AAZ26733	Consensus sequence
28	422.8	88.1	580	20	AAZ26728	Consensus sequence
29	422.8	88.1	580	20	AAZ00401	Hepatitis C virus
30	422.8	88.1	1249	16	AAQ79772	Hepatitis C virus
31	422.8	88.1	1249	20	AAZ26739	Consensus sequence
32	422.8	88.1	1270	19	AAV60668	Fragment #1 isolat
33	422.8	88.1	3401	15	AAQ64069	Non-A, non-B hepat
34	422.8	88.1	3401	16	AAT30387	5'UTR/CORE/ENV/NS1
35	422.8	88.1	3461	15	AAQ64068	Non-A, non-B hepat
36	422.8	88.1	3461	16	AAT30386	5'UTR/CORE/ENV/NS1
37	422.2	88.0	642	17	AAT12704	HCV E1 construct H
38	422.2	88.0	642	24	AA148913	Hepatitis C virus
39	421.6	87.8	1882	13	AAQ27159	NANB hepatitis vir
40	421.2	87.8	577	14	AAQ35086	HCV envelope regio
41	421.2	87.8	580	16	AAQ79758	Hepatitis C virus
42	421.2	87.8	580	16	AAQ79759	Hepatitis C virus
43	421.2	87.8	580	16	AAQ79760	Hepatitis C virus
44	421.2	87.8	795	17	AAQ12705	HCV E1 construct H
45	421.2	87.8	795	24	AA148914	Hepatitis C virus

ALIGNMENTS

RESULT 1
AAT12708
ID AAT12708 standard; DNA; 480 BP.

XX
AC AAT12708;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC113A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 483 BP; 85 A; 152 C; 123 G; 123 T; 0 other;

Query Match 98.8%; Score 474.2; DB 24; Length 483;
Best Local Similarity 99.4%; Pred. No. 9.4e-129; Indels 0; Gaps 0;
Matches 476; Conservative 0; Mismatches 3;
QY 1 ATCTCGGTTGCTCTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGACCATACCA 60
DB 1 ATGCGCGTTGCTCTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGACCATACCA 60
QY 61 GCTTCGCTTATGAAGTGCACACGTTGCGGGGTGTACATGTACAGCACTGCTCC 120
DB 61 GCTTCGCTTATGAAGTGCACACGTTGCGGGGTGTACATGTACAGCACTGCTCC 120
QY 121 AACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTCGCTGCC 180
DB 121 AACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTCGCTGCC 180
QY 181 TGCCTTCGGAGGCAACTCTCTCCGTTGTGGGTGGCGTCACTCCACGCTCGCGGC 240
DB 181 TGCCTTCGGAGGCAACTCTCTCCGTTGTGGGTGGCGTCACTCCACGCTCGCGGC 240
QY 241 AGGAACGCGAGCTCCCAACAGCAATACGAGCGCACCTCGATTTGCTGTTGGGCT 300
DB 241 AGGAACGCGAGCTCCCAACAGCAATACGAGCGCACCTCGATTTGCTGTTGGGCT 300
QY 301 GCTGCTTCTCTCCGTTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 360
DB 301 GCTGCTTCTCTCCGTTATGATGAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 360
QY 361 CAGCTGTTTCACTTCTCACTTCGCGGATCAAAACAGTACAGGCTGCACTGCTCAATC 420
DB 361 CAGCTGTTTCACTTCTCACTTCGCGGATCAAAACAGTACAGGCTGCACTGCTCAATC 420
QY 421 TATCCCGCCATGTATCAGGTCACCGCATGCTTGGGATATGATGATGATGATGATG 479
DB 421 TATCCCGCCATGTATCAGGTCACCGCATGCTTGGGATATGATGATGATGATGATG 479

RESULT 5
AAQ24467
ID AAQ24467 standard; DNA; 1880 BP.
XX
XX AAQ24467;
XX
XX 09-NOV-1992 (first entry)
XX
XX NANB hepatitis virus strain HC-J4 genome.
DE
XX non-A, non-B hepatitis virus; NANBH; PCR; amplification
KW polymerase chain reaction; vaccine; antibody; ss.
XX
XX Non-A, non-B hepatitis virus.
XX

Key Location/Qualifiers
CDS 342..1880
FT /*tag= a
FT FT /*label= HC-J4
XX
XX EP485209-A.
XX
XX 13-MAY-1992.
XX
XX 07-NOV-1991; 91EP-0310297.
XX

PR 08-NOV-1990; 90JP-0304405.
XX (IMMO) IMMUNO JAPAN INC.
XX
XX Nakamura T, Okamoto H;
XX WPI; 1992-160959/20.
DR P-PSDB; AAR24087.
XX
XX Recombinant cDNA of NANBH virus strain HC-J5 and corresp.
PT peptides - useful for diagnosis and in vaccines and immunological
PT pharmaceuticals
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX This sequence is the genome of the non-A, non-B hepatitis virus
CC (NANBH) strain HC-J4. This sequence was derived by amplification
CC by polymerase chain reaction. The nucleotide sequences derived from
CC this amplification can be used to detect NANBH infection which could
CC not be detected by conventional methods. The detection kits allow
CC highly specific and sensitive detection at an early phase of
CC infection. The polypeptide product of this coding sequence can be used
CC for the manufacture of vaccines and immunological pharmaceuticals
CC and also to produce antibodies specific to NANBH.
XX
XX Sequence 1880 BP; 333 A; 593 C; 556 G; 398 T; 0 other;
SQ

Query Match 90.2%; Score 432.8; DB 13; Length 1880;
Best Local Similarity 94.3%; Pred. No. 2e-116;
Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 2 TGTCCGGTTGCTCTTCTCTATCTTCTTGGCCCTGCTGCTCTGACCATACCA 61
DB 847 TGCCTTCGTTGCTCTTCTCTATCTTCTTGGCTTGTCTGCTGCTGCTGCTG 906
QY 62 CTTCCGCTTATGAAGTGCACACGTTCCGGGTGTACATGTACAGCACTGCTCCA 121
DB 907 CTTCCGCTTATGAAGTGCACACGTTCCGGGTGTACATGTACAGCACTGCTCCA 966
QY 122 ACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTCGCTCC 181
DB 967 ACTCAAGCATGTATGATGAGGAGCGGACATCATGCACACCCCGGTCGCTCC 1026
QY 182 GGTTCGGAGGCACTCTCCGTTGCTGGGTGCGCTCACTCCACGCTCGCGGCCA 241
DB 1027 GGTTCGGAGGCACTCACTCCGTTGCTGGGTGCGCTCACTCCACGCTCGCGGCCA 1086
QY 242 GGAACCCAGCGTCCCAACAGCAATACGACGCCACGTCGATTTGCTGTTGGGCTG 301
DB 1087 GGAATGCCAGGCTCCCACTACGACAAATACGACGCCACGTCGATTTGTTGGGCGG 1146
QY 302 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 361
DB 1147 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTGTTC 1206
QY 362 AGCTGTTCACTTCTCACTTCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATCT 421
DB 1207 AGCTGTTCACTTCTCGCTTCGCGGCATGACAGACTGCGAGGACTGCAACTGCTCAATCT 1266
QY 422 ATCCCGCCATGTATCAGGTCACCGCATGCTTGGGATATGATGATGATGATGATG 477
DB 1267 ATCCCGCCATGTATCAGGTCACCGCATGCTTGGGATATGATGATGATGATGATG 1322

RESULT 6
ABA03491
ID ABA03491 standard; DNA; 2187 BP.
XX
XX ABA03491;
XX
XX 15-MAR-2002 (first entry)
XX
XX Cuticle protein 1 and 2 secreting hepatitis C virus related DNA #1.

XX Cuticle protein 1; cuticle protein 2; hepatitis C virus; ds.
XX Unidentified.
XX Key Location/Qualifiers
XX CDS 1..2187
XX /*tag= a
XX /*product= "AAM47264"
XX /*partial
XX /*note= "no stop codon"
XX
XX KR97065713-A.
XX 13-OCT-1997.
XX 19-MAR-1996; 96KR-0007404.
XX 19-MAR-1996; 96KR-0007404.
XX (GLDS) LG CHEM LTD.
XX Choo SH, Lee IH, Ryoo WS;
XX WPI; 1998-492654/42.
XX P-PSDB; AAM47264.
XX
XX Cuticle protein 1 and 2 secreting hepatitis C virus (Japanese)
XX NoAbstract -
XX
XX Disclosure; Page 2-4; 7pp; Korean.
XX
XX The present invention relates to cuticle protein 1 and 2 secreting
XX hepatitis C virus. The present sequence is a coding sequence
XX provided in the exemplification of the invention.
XX
XX Sequence 2187 BP; 406 A; 669 C; 631 G; 481 T; 0 other;
XX
XX Query Match 90.2%; Score 432.8; DB 19; Length 2187;
XX Best Local Similarity 94.3%; Pred. No. 2.1e-116;
XX Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
XX
QY 2 TGTCCGGTGTCTTTCTCTATCTTCTCTATCTTCTCTGTGGCCCTGTCTGTCTGACCATACCAG 61
DB 506 TGCCTGGTGTCTTTCTCTATCTTCTCTATCTTCTCTGTGGCTTGTCTGTCTGTTGACCATCCAG 565
QY 62 CTTCCGCTTATGAAGTCCCAAGTGTCTCGGGTGTACCATGTCCAGAACGACTGTCTCCA 121
DB 566 CTTCCGCTTATGAAGTCCCAAGTGTCTCGGGTGTACCATGTCCAGAACGACTGTCTCCA 625
QY 122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGTGTGCGCT 181
DB 626 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCATATCTCCGGTGTGCGCT 685
QY 182 GCCTTCGGAGGCAACTCTCCGTTGTGGTGTGGTGTCTCTCCACGCTCGCGCCA 241
DB 686 GCCTTCGGAGGCAACTCTCCGTTGTGGTGTGGTGTCTCTCCACGCTCGCGCCA 745
QY 242 GGAACGCGAGCTCCCAACACATACGACGCGCATGATGCTGATGCTGTTGGGCTG 301
DB 746 GGAATGCCAGCTCCCAACATACGACGCGCATGATGCTGATGCTGTTGGGCTG 805
QY 302 CTGCTTCTCTGCGGTATGATGAGGAGTCTCTGCGGATCTGTTTCTCTGTTTCCC 361
DB 806 CTGCTTCTCTGCGGTATGATGAGGAGTCTCTGCGGATCTGTTTCTCTGTTTCCC 865
QY 362 AGCTGTTCACCTTCTCACTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT 421
DB 866 AGCTGTTCACCTTCTCACTCGCGGATCAAAACAGTACAGGACTGCAACTGCTCAATCT 925
QY 422 ATCCCGGCCATGATACAGGTACCGGATGCGCTTGGGATATGATGATGAAGTGTAA 477
DB 926 ATCCCGGCCATGATACAGGTACCGGATGCGCTTGGGATATGATGATGAAGTGTAA 981

RESULT 7
AAQ43889
ID AAO43889 standard; cDNA to mRNA; 2540 BP.
XX
XX AAO43889;
XX
XX 21-OCT-1993 (first entry)
XX
XX NANB hepatitis virus polynucleotide N-2540-2.
XX
XX Non-A, non-B; virus; polymerase chain reaction; detection;
XX sensitive; specific; HCV; NANBH; ss.
XX
XX Non-A, non-B hepatitis virus.
XX
XX Key Location/Qualifiers
XX CDS 342..2540
XX /*tag= a
XX 5'UTR 1..341
XX /*tag= b
XX /*note= "from 5' terminal of NANBH virus RNA"
XX
XX JP05091884-A.
XX
XX 16-APR-1993.
XX
XX 10-APR-1991; 91JP-0196175.
XX
XX 12-JUN-1990; 90JP-0153401.
XX 08-NOV-1990; 90JP-0304405.
XX
XX (NAKA/) NAKAMURA T.
XX
XX WPI; 1993-199637/25.
XX P-PSDB; AAR38279.
XX
XX Antigen related to non-A and non-B hepatitis virus - comprises
XX non-translation region comprising 340 - 341 mols. of nucleotides,
XX non-translation region comprising 1885 - 2551 mols. of
XX nucleotides including region 1,149 and, etc.
XX
XX Claim 3; Page 19-20; 73pp; Japanese.
XX
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
XX The polypeptide it encodes may be used in a system for detecting
XX NANB hepatitis. This method is highly specific and sensitive, and
XX can detect NANB hepatitis virus which could not be detected by
XX conventional methods.
XX
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
XX
XX Query Match 90.2%; Score 432.8; DB 14; Length 2540;
XX Best Local Similarity 94.3%; Pred. No. 2.2e-116;
XX Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
XX
QY 2 TGTCCGGTGTCTTTCTCTATCTTCTCTATCTTCTCTGTGGCCCTGTCTGTCTGACCATACCAG 61
DB 847 TGCCTGGTGTCTTTCTCTATCTTCTCTATCTTCTCTGTGGCTTGTCTGTCTGTTGACCATCCAG 906
QY 62 CTTCCGCTTATGAAGTCCCAAGTGTCTCGGGTGTACCATGTCCAGAACGACTGTCTCCA 121
DB 907 CTTCCGCTTATGAAGTCCCAAGTGTCTCGGGTGTACCATGTCCAGAACGACTGTCTCCA 966
QY 122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGTGTGCGCT 181
DB 967 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCATATCTCCGGTGTGCGCT 1026
QY 182 GCCTTCGGAGGCAACTCTCCGTTGTGGTGTGGTGTCTCTCCACGCTCGCGCCA 241
DB 1027 GCCTTCGGAGGCAACTCTCCGTTGTGGTGTGGTGTCTCTCCACGCTCGCGCCA 1086

Db 967 ACTACAGCATTTGTATGAGGAGGACATGATCATGCACTATCCCGGGTGGTGCCT 1026
Qy 182 GCGTTCCGGAGGGGAACTCCCTCCGTTGGTGGGGCTCACTCCACGCTCGGGCCA 241
Db 1027 GCGTTCCGGAGGACAAACAGCTCCCGTTGGTGGGTAGCGCTCACTCCACGCTCGGGCCA 1086
Qy 242 GGAACGCCAGGTCCTCCACACGACAAATACGAGCCACGTCGATTTGCTCGTTGGGGCTG 301
Db 1087 GGAATGCCAGGTCCTCCACACTACGACAAATACGAGCCACGTCGATTTGCTCGTTGGGGCTG 1146
Qy 302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361
Db 1147 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1206
Qy 362 AGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db 1207 AGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 1266
Qy 422 ATCCGCGCATGTATCAGGTCAACGCGCATGGCTTGGGATATGATGATGAATGGTAA 477
Db 1267 ATCCGCGCATTTATCAGGTCAACGCGCATGGCTTGGGATATGATGATGAATGGTCA 1322

RESULT 9

AAQ15363
ID AAQ15363 standard; DNA; 1863 BP.

XX AC AAQ15363;

XX XX Polymerase chain reaction; PCR; primer: amplify; detection; NANBH;

DT 25-MAR-2003 (updated)

DT 17-MAR-1992 (first entry)

XX DE Fragment of NANB hepatitis virus strain HC-J4.

XX XX Post-transfusion; non-A, non-B hepatitis; HCV; detection primer; PCR;

KW KW polymerase chain reaction; ss.

XX OS Hepatitis C virus.

XX XX Key Location/Qualifiers

FT 1..324

FT 5'UTR

FT /*tag= a

FT CDS 325..1863

FT /*tag= b

FT EP461863-A.

XX PN

XX XX 18-DEC-1991.

XX PF 11-JUN-1991; 91EP-0305270.

XX PF 12-JUN-1990; 90JP-0153402.

XX PA (IMMO) IMMUNO JAPAN INC.

XX PI Okamoto H, Ogikubo Y, Nakamura T;

XX WPI; 1991-370834/51.

XX PT Oligo-nucleotide primers - derived from and used to detect and

XX PT diagnose non-A, non-B hepatitis virus

XX PS Claim 1; Page 7; 23pp; English.

XX XX NANB hepatitis virus strain HC-J4 was isolated from a plasma sample

XX CC of a chimpanzee challenged with NANB hepatitis for infectivity but

XX CC which tested negative for HCV antibody by Ortho HCV Ab ELISA test.

XX CC RNA was isolated from the sample and reverse transcribed into cDNA.

XX CC The 513 amino acids encoded by the CDS were determined but are not

XX CC given in the specification (and hence are not included in A-Geneseq).

XX CC A study of the deduced sequence suggested that the CDS encodes NANBH

XX CC virus core proteins. Primers for detecting NANB hepatitis virus were

Qy 242 GGAACGCCAGGTCCTCCACACGACAAATACGAGCCACGTCGATTTGCTGGGGCTG 301
Db 1087 GGAATGCCAGGTCCTCCACACTACGACAAATACGAGCCACGTCGATTTGCTGGGGCTG 1146
Qy 302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 361
Db 1147 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTCGGATCTGTTTCTGTTTCCC 1206
Qy 362 AGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db 1207 AGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATCT 1266
Qy 422 ATCCGCGCATGTATCAGGTCAACGCGCATGGCTTGGGATATGATGATGAATGGTAA 477
Db 1267 ATCCGCGCATTTATCAGGTCAACGCGCATGGCTTGGGATATGATGATGAATGGTCA 1322

RESULT 8

AAQ63753
ID AAQ63753 standard; cDNA to mRNA; 2540 BP.

XX AC AAQ63753;

XX XX 30-JAN-1995 (first entry)

XX DE NANBH genomic fragment #2.

XX KW Polymerase chain reaction; PCR; primer: amplify; detection; NANBH;

XX KW non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.

XX OS Synthetic.

XX PN JP06125777-A.

XX PD 10-MAY-1994.

XX PF 20-JUN-1991; 91JP-0247120.

XX PF 20-JUN-1991; 91JP-0247120.

XX PR (NAKA/) NAKAMURA T.

XX PA WPI; 1994-187937/23.

XX DR Oligonucleotide primer pairs specific for non-A, non-B hepatitis

XX PT virus - used for high sensitivity detection of non-A non-B (NANB)

XX PT hepatitis virus

XX PS Disclosure; Page 24-25; 25pp; Japanese.

XX XX The sequences given in AAQ63752-53 represent fragments of the non-A,

XX CC non-B hepatitis virus (NANBH) genome. These fragments were amplified

XX CC using the primers given in AAQ63732-51. These primers were used in the

XX CC detection of NANBH. The primers are based on the 5'-terminal region and

XX CC the core protein coding region. The method allows highly sensitive

XX CC detection of NANBH.

XX SQ Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

XX Query Match 90.2%; Score 432.8; DB 15; Length 2540;

XX Best Local Similarity 94.3%; Pred. No. 2.2e-116;

XX Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 2 TGTCCGGTGTCTTTCTATCTCTATCTCTCTTGGCCCTGCTCTGCTGACCATACAG 61

Db 847 TGCCTGGTGTCTTTCTATCTCTCTTGGCTTGTCTGCTGCTGACCATACAG 906

Qy 62 CTTCCGCTTATCAAGTGGCAACGTCGCGGGGTGTACCATGTCAAGCAACGACTGCTCCA 121

Db 907 CTTCCGCTTATCAAGTGGCAACGTCGCGGGGTGTACCATGTCAAGCAACGACTGCTCCA 966

Qy 122 ACTCAAGCATAGTGTATGAGGAGCGGACATGATGATGACACCCCGGGTGGTGCCT 181

	Key	Location/Qualifiers
H	CDS	342..9374
T		/*tag= a
I		
X	WO9904008-A2.	
N		
N		
X		
X	28-JAN-1999.	
D		
D		
X		
X	16-JUL-1998;	98WO-US14688.
P		
X		
X	27-JAN-1998;	98US-0014416.
R		
R	18-JUL-1997;	97US-0053062.

X
I Bukh J, Emerson SU, Purcell RH, Yanagi M;
R WPI; 1998-132252/11.
R P-PSDE; AAW98022.
X
X
X
T
T
T
T
X
X
X
X

New isolated hepatitis C virus nucleic acids - used to develop products for the diagnosis, prevention and treatment of HCV infections and for developing screening assays

Claime 3; Fig 14A-F; 126pp; English.

genome of infectious hepatitis C virus (HCV) genotype 1b strain HC-J4 (ATCC 209596) that is capable of expressing this virus when transfected into cells. HC-J4 was obtained from acute phase plasma of a chimpanzee experimentally infected with serum containing HC-J4/91. The claimed infectious nucleic acid sequence can be used to produce chimeric genomes (see AAX24833) consisting of the open reading frames of infectious nucleic acid sequences of other HCV genotypes (see AAX24833) (such as the HCV genotype 1b strain

2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to the introduction of mutations or deletions into infectious nucleic acid sequences in order to produce an attenuated HCV virus suitable for vaccine development. Infectious nucleic acid sequences can also be used to produce attenuated virus via passage *in vitro* or *in vivo* of the viruses produced by transfection of a host cell with the infectious nucleic acid sequence. Vaccines comprising one or more of the nucleic acid sequences of the invention are also provided.

used to immunise mammals, especially humans, against hepatitis C. The nucleic acid sequences can also be used to induce protective immunity against the virus. The nucleic acid sequences or their encoded proteases (e.g. NS3 protease) can additionally be used to develop screening assays to identify antiviral agents for HCV.

Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 other;

Query Match	89.2%;	Score 428;	DB 20;	Length 5553;
Best Local Similarity	93.7%;	Pred. No. 9.1e-115;		
Matches 446;	Conservative	0;	Mismatches 30;	Indels 0;
	Gaps	0;		

QY	2	TGTCGGGTGCCTTTCTCTATCTCTCTTGGCCCTGTGTGCTCTGTACCATACCAAG	61
Db	847	TGCCCGGTTGCTCTTTCTCTATCTCTCTTGGCTCTGCTGTCCTGTTGACCATCCCAAG	906
QY	62	CTTCCGCTTATGAAGTGCACAGCTGTCCGGGGGTACCATGTACAGAACGACTGTCTCA	121
Db	907	CTTCCGCTTATGAAGTGCACAGCTGTCCGGGATATACCATGTACAGAACGACTGTCTCA	966
QY	122	ACTCAAGCATPAGTGTATGAGCGACGAGCATGATCATGCAACCCCGGGTGGTGGCCCT	181
Db	967	ACTCAAGCATTGTGTATGAGCGACGAGCATGATCATGCATCTCCCGGGTGGTGGCCCT	1026
QY	182	CGCTTCGGAGGGCAATCTCCCGTGTCTGGGTGGCGCTCATCTCCACACGCTCGCGSCCA	241
Db	1027	GTGTTTCAGGAGGGTAAACAGTCCCGTTGCTGGGTAGCGCTCATCTCCACACGCGCGCGCCA	1086

QY 242 GGAAAGCCAGCGTCCCCCAACAGACAAATACAGACGCCACGTGCGATTTTGCTGTTGGGGCTG 301

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1234.83 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-11
Perfect score: 480
Sequence: 1 ATGTCGGTGTCTTCTC.....TGATGATGAACCTGGTAATAG 480

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	65.8	13.7	488	9	AV755731
C 2	54.6	11.4	492	9	AV758366
C 3	43.2	9.0	534	14	CD040840
C 4	41.6	8.7	664	29	BZ645446

C	5	40.6	8.5	526	9	AL825643
6	7	40.2	8.4	399	9	AV638521
8	4	40.2	8.4	434	9	AV637507
9	4	40.2	8.4	440	9	AV637983
10	4	40.2	8.4	450	9	AV637259
11	4	40.2	8.4	451	9	AV637328
12	4	40.2	8.4	451	9	AV637643
13	4	40.2	8.4	453	9	AV634724
14	4	40.2	8.4	454	9	AV637050
15	4	40.2	8.4	456	9	AV635382
16	4	40.2	8.4	473	9	AV632765
17	4	40.2	8.4	481	9	AV635503
18	4	40.2	8.4	485	9	AV632811
19	4	40.2	8.4	506	9	AV392445
20	4	40.2	8.4	508	9	AV634095
21	4	40.2	8.4	526	9	AV641895
22	4	40.2	8.4	533	9	AV638125
23	4	40.2	8.4	537	9	AV632335
24	4	40.2	8.4	588	9	AV387329
25	4	40.2	8.3	1186	13	BX421743
26	39.4	39.2	8.2	1201	9	AL565958
27	39.2	39.2	8.2	624	14	CD206870
28	39	39	8.1	656	14	CB924688
29	39	39	8.1	497	9	AV633658
30	39	39	8.1	610	14	CB657655
31	39	39	8.1	856	29	BZ578381
32	38.4	38.4	8.0	872	29	BZ555011
33	38	38	7.9	645	29	CNS01213
34	37.8	37.8	7.9	771	29	BZ530934
35	37.8	37.8	7.9	490	9	AV634529
36	37.8	37.8	7.9	705	14	CA618797
37	37.6	37.6	7.8	1039	13	BX415186
38	37.6	37.6	7.8	309	12	BI098866
39	37.4	37.4	7.8	431	9	AV639153
40	37.4	37.4	7.8	394	12	BZ209795
41	37.2	37.2	7.8	501	9	AV638474
42	37.2	37.2	7.8	431	9	AV641448
43	37	37	7.7	792	13	BX391120
44	36.8	36.8	7.7	910	29	CNS0060N
45	36.8	36.8	7.7	214	12	BM868105
				913	14	CA487901

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMAFB03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

```
source
1..488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN
Query Match 13.7%; Score 65.8; DB 9; Length 488;
Best Local Similarity 67.0%; Pred. No. 2.3e-07;
Matches 124; Conservative 0; Mismatches 57; Indels 4; Gaps 2;

QY 292 GTTGGGCTGCTGCTTCTGTTCCGTATGTACGTGGGGATCTCTGGGATCTGTTTC 351
Db 472 GTGGTGTCACACTCGCTCTCTAGCTGTGGGACCTCTGGCAAGAGTGATG 413
QY 352 CTTGTTTCCAGCTGTTCACCTTCTCAGCTCGCGGCATCAACAGTACAGACTGCAAC 411
Db 412 CTTGCAGTTTCAGCTGATCA---TCTGGCTTCAGCACCATGAGTTTGTGATGATGCAAC 356
QY 412 TGCTCAATCTATCCGGCCATGTATCAGTCAACCGCATG-GTTGGGATATGATGAA 470
Db 355 TGCTCCATCTATCTGGGCCATCAGTGGACCGTATGAGCATGGGACATGATGAA 296
QY 471 CTGCT 475
Db 295 CTGCT 291

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366.1 GI:10916214
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Gu, J., Zhao, M., Qian, B., Liu, F., Ou, J., Gao, X., Cheng, Z., Xu, Z., Zeng
AUTHORS L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN
```

```
Query Match 11.4%; Score 54.6; DB 9; Length 492;
Best Local Similarity 65.7%; Pred. No. 0.00027;
Matches 111; Conservative 0; Mismatches 54; Indels 4; Gaps 2;

QY 308 TCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTCTGTTTCTGTTTCCAGCTGT 367
Db 457 TGTATCAGCTCACTACGTGTTGGACCTCTGGCTTGGGGTATCGTTGAGCCCACTGA 398
QY 368 TCACCTTCTCACTCGCGGCATCAACAGTACAGACTGCAACTGCTCAATCTATCCCG 427
Db 397 TTA---TCTCTCAGCAGCAACATTTGGTTTGTGCAAGATGCAACTGCTATCTATCTG 341
QY 428 GCCATGTATCAGTTC-ACCGCATGGCTTGGGATATGATGATGAACTGGT 475
Db 340 GCTGCATCACTGGACATACAGTATGGCATAGCTATGATGATGAACTGGT 292

RESULT 3
CD040840 534 bp mRNA linear EST 09-MAY-2003
LOCUS pHB036xB09f_300663 psHB: Infected hypocotyl soybean host. 48 hrs
DEFINITION post infection Phytophthora sojae cDNA clone sHB036B09 5, mRNA
sequence.
ACCESSION CD040840
VERSION CD040840.1 GI:30502701
KEYWORDS EST.
Phytophthora sojae
Phytophthora sojae
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
ORGANISM Phytophthora.
REFERENCE 1 (bases 1 to 534)
AUTHORS Tyler, B.M., Judelson, H.S., Gijzen, M., Dean, R.A. and Waugh, M.E.
TITLE USDA-IPAFS: Expression of Phytophthora sojae genes during infection
and propagation
JOURNAL Unpublished
COMMENT Contact: Tyler B
Tyler lab
VBI
1880 Pratt Dr., Blacksburg, VA 24061, USA
Tel: 540-231-7318
Email: bmtyle@vt.edu
PCR Primers
FORWARD: BK reverse
Plate: 036 row: B column: 09
Seq primer: BK reverse
High quality sequence stop: 534.
Location/Qualifiers
1..534
/organism="Phytophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="sHB036B09"
/tissue_type="infected host tissue"
/cell_line="p6497"
/dev_stage="48 hour post infection"
/clone_lib="psHB: Infected hypocotyl soybean host. 48 hrs
post infection"
/note="Vector: psB-CMV; Site_1: EcoRI; Site_2: XhoI;
USDA-IPAFS: Expression of Phytophthora sojae genes during
infection and propagation."
BASE COUNT 101 a 187 c 159 g 87 t
ORIGIN
Query Match 9.0%; Score 43.2; DB 14; Length 534;
Best Local Similarity 47.4%; Pred. No. 0.36;
Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGGCAACGTGTCGGGGTGTACCATGTACGACGACGCTCTCAACTCAAGC 129
Db 200 TACGGCTGGCGGAGATTAGGTATCCGATGGCTTCGCCGCTTCTACAACTGGACC 259
QY 130 ATAGTGTATGAGGAGCGGACATGATCATGCACACCCCCGGTGGCTCGGTTTCGG 189
```



```

Db      260 TCATGGACAGAGAGGCCCCGATCATGCTGACCCCAAGACGGTGCCCAACATTCAC 319
Qy      190 GAGGGCAACTCTCCGTTGCTGGGTGGGGCTCACTCCCAACGCTCGCGGCCAGGAACGGC 249
Db      320 CACTACGGCGGCACCATCTCGGCTCGAACCGTGGTGGCTTCGACGTGGACAGATTATC 379
Qy      250 AGGTTCCCAACAGCAATACAGCCACGCTGATTTGCTGTTGGGGCTGCTGCTTTC 309
Db      380 AACTTCTTACGACGACGACGCGCTCTCGAGGTATACGTATCGTCGCGGTGACGGCACCCAC 439
Qy      310 TGTTCGCTATGTACGTGGGGATCTCTGGG 341
Db      440 CGTGCCGCCACAGATCTCGGAGGAGTCCG 471

RESULT 4
LOCUS   BZ645446/c
DEFINITION BZ645446 664 bp DNA linear GSS 29-JAN-2003
ACCESSION BZ645446
VERSION   BZ645446.1 GI:28107610
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
           A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
           R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE     Consortium for Maize Genomics
JOURNAL   Unpublished
COMMENT   Other GSSs: OGCBJ86TM
           Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source    1..664
           /organism="Zea mays"
           /mol_type="genomic DNA"
           /strain="B73"
           /db_xref="taxon:4577"
           /clone="ZMBMa0133P04"
           /clone_lib="ZM 0.7 1.5 KB"
           /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
           methylation filtered genomic DNA library"

BASE COUNT  103 a 238 c 222 g 101 t
ORIGIN
Query Match      8.7%; Score 41.6; DB 29; Length 664;
Best Local Similarity 50.5%; Pred. No. 1.1;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy      119 CCAACTCAAGCATAGTGTATGAGCGACGCGACATGATCATGACACCCCGGGTGGGTGC 178
Db      266 CCGACGACGACACATCGTGGTGGCGCCCAAGACGAGGTGCGGACGAGGACCGGGAC 207
Qy      179 CTGCGTTGGGAGGGCACTCTCCGTTGCTGGGTGGCGCTCATCTCCACGCTCGCGG 238
Db      206 CGTACCTGGGTGGCGAGCTCCACGAAGTGTAGAAGCCCTGCGCTGCTGTGGCGG 147
Qy      239 CCAGGAACCGCGTCCCAACAGCAATACGACGCCACGTCGATTCTCTGTTGGGG 298
Db      146 ACAGGAAGGCCAGCGTCTCCACGACGAGCCACCACGAGCGGTACCCCTTCTCGGTGGT 87

```

```

Qy      299 CTGCTGCTTTCTGTTCCGCT 318
Db      86 GTGGTGGTGGTGGTGGTGGT 67

RESULT 5
LOCUS   AL825643/c
DEFINITION AL825643 p.234 Triticum aestivum cDNA clone A09_p234_plate_14, mRNA
ACCESSION AL825643
VERSION   AL825643.1 GI:21837164
KEYWORDS  EST.
SOURCE    Triticum aestivum (bread wheat)
ORGANISM  Triticum aestivum
REFERENCE
AUTHORS   Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
           Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
           Edwards,K.
TITLE     A BSRRC-funded wheat EST resource for the academic community
JOURNAL   Unpublished
COMMENT   Contact: Barker G
           Institute of Arable Crop Research
           Long Ashton, Bristol BS41 9AF United Kingdom.

FEATURES
source    1..526
           /organism="Triticum aestivum"
           /mol_type="mRNA"
           /cultivar="mercia"
           /db_xref="taxon:4565"
           /clone="A09_p234_plate_14"
           /tissue_type="drought stressed seedlings"
           /dev_stage="21 days old"
           /clone_lib="p:234"

BASE COUNT  97 a 179 c 164 g 86 t
ORIGIN
Query Match      8.5%; Score 40.6; DB 9; Length 526;
Best Local Similarity 57.5%; Pred. No. 1.8;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy      150 CATGATCATGCACACCCCGGGTGGTGGCTGCTCGCTCGGGAGGCAACTCTCCCGTTG 209
Db      151 CTCTCTCCCGAACCCCGGGTGGATCAGTCGCGAGTTCTTCGGCAGCCCTTCTTGCC 92
Qy      210 CTGGGTGGCGCTCACTCCACGCTCGCGCCAGGAACGCCAGCGTCCCAACACGACAAT 269
Db      91 CTGGGCGGCGCGCGGACGCGCTCGGCACAGGAACACAGCACCGCGCGCGCGGAC 32
Qy      270 ACGACGC 276
Db      31 CACACCC 25

RESULT 6
LOCUS   AV638521
DEFINITION AV638521 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
ACCESSION AV638521
VERSION   AV638521.1 GI:10781841
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
REFERENCE
AUTHORS   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
           Chlamydomonadaceae; Chlamydomonas.
           1 (bases 1 to 399)
           Asanuma,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
           Nakamura,Y. and Tabata,S.
           Generation of expressed sequence tags from low-CO2 and high-CO2

```

Mon Dec 22 13:28:34 2003

```

adapted cells of Chlamydomonas reinhardtii
DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
    source
        1..399
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="C9"
            /db_xref="taxon:3055"
            /clone_lib="HC087d07_r"
            /clone_lib="Chlamydomonas reinhardtii 5% CO2"
            /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
            xhoI; The cDNA library was constructed from cells cultured
            in a medium with bubbling air containing 5% carbon
            dioxide"
BASE COUNT 70 a 144 c 122 g 63 t
ORIGIN
Query Match 8.4%; Score 40.2; DB 9; Length 399;
Best Local Similarity 51.4%; Pred. No. 2.2;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 30 CTTGGCCCTGTCCTGTCGACCATACAGCTTCCGCTTATGAAGTGGCAACGTGTC 89
DB 44 CTTGACCCCTGACCGCTGCTCATCTGTCGACCACTTCCGCTGGCACCATCTT 103
QY 90 CGGGGTGTACATGTACGAACTGCTCCTCAACTCAAGCATAGTGTATGAGGAGCGGA 149
DB 104 CGGGTGTGGCCGGGACAGCTGACCAATCACCGCGCGAGCAGGTGGTGGCGG 163
QY 150 CATGATCATGCACACCCCGGTCGCTGCTGCGTTCGGGAGGCACTCTCCCGTTG 209
DB 164 CATGGCATCTACGTCCTCCGCGCTGTTCTGCAATTCCTGAAGGACGCCCGGCTG 223
QY 210 C 210
DB 224 C 224

RESULT 7
AV637507 434 bp mRNA linear EST 15-DEC-2000
LOCUS AV637507 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION CDNA clone HC074d01_r 5', mRNA sequence.
ACCESSION AV637507.1 GI:10780827
VERSION AV637507
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 434)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
    source
        1..434
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="C9"
            /db_xref="taxon:3055"
            /clone_lib="HC080c04_r"
            /clone_lib="Chlamydomonas reinhardtii 5% CO2"
            /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
            xhoI; The cDNA library was constructed from cells cultured
            in a medium with bubbling air containing 5% carbon
            dioxide"
BASE COUNT 76 a 156 c 137 g 71 t
ORIGIN
Query Match 8.4%; Score 40.2; DB 9; Length 434;
Best Local Similarity 51.4%; Pred. No. 2.2;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 30 CTTGGCCCTGTCCTGTCGACCATACAGCTTCCGCTTATGAAGTGGCAACGTGTC 89
DB 21 CTTGACCCCTGACCGCTGCTCATCTGTCGACCACTTCCGCTGGCACCATCTT 80
QY 90 CGGGGTGTACATGTACGAACTGCTCCTCAACTCAAGCATAGTGTATGAGGAGCGGA 149
DB 81 CGGGTGTGGCCGGGACAGCTGACCAATCACCGCGCGAGCAGGTGGTGGCGG 140
QY 150 CATGATCATGCACACCCCGGTCGCTGCTGCGTTCGGGAGGCACTCTCCCGTTG 209
DB 141 CATGGCATCTACGTCCTCCGCGCTGTTCTGCAATTCCTGAAGGACGCCCGGCTG 200
QY 210 C 210
DB 201 C 201

RESULT 8
AV637983 440 bp mRNA linear EST 15-DEC-2000
LOCUS AV637983 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION CDNA clone HC080c04_r 5', mRNA sequence.
ACCESSION AV637983
VERSION AV637983.1 GI:10781303
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 440)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
    source
        1..440
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="C9"
            /db_xref="taxon:3055"
            /clone_lib="HC080c04_r"
            /clone_lib="Chlamydomonas reinhardtii 5% CO2"
            /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
            xhoI; The cDNA library was constructed from cells cultured
            in a medium with bubbling air containing 5% carbon
            dioxide"
BASE COUNT 76 a 156 c 137 g 71 t
ORIGIN

```

```

Query Match      8.4%; Score 40.2; DB 9; Length 440;
Best Local Similarity 51.4%; Pred. No. 2.2; Mismatches 0; Gaps 0;
Matches 93; Conservative 0; Indels 88;

QY 30 CTTGGCCCTGCTGCTCTGCTGACCATACAGCTTCCGCTTATGAAGTGGCAACGTGTC 89
DB 80 CTTGACCCCTGGAGGGCTGCTCCATCGTGACACCACTTCCGCGTGGACCATCTT 139

QY 90 CCGGGTGTTACCATGTTCACGAAGCTGCTCCCACTCAAGCATAGTGTATGAGGACGGGA 149
DB 140 CGCGCTGTGGCCGGCGGACAAAGTGACCAACATCACCGCGCGGAGCAGTGGCTGCCGG 199

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGGGCAACTCCTCCCGTTG 209
DB 200 CATGGCATCTAGGTCCCGCACCGCTGTTCTGTCATTGCCCTGAAGGACGCCCCGGCTG 259

QY 210 C 210
DB 260 C 260

RESULT 9
AV637259
LOCUS
DEFINITION AV637259 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HC070906_r 5', mRNA sequence.
ACCESSION AV637259
VERSION
KEYWORDS
SOURCE
ORGANISM Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 450)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..450
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC070906_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT 74 a 153 c 148 g 75 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 450;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTGCTCTGCTGACCATACAGTTCGCTTATGAAGTGGCAACGTGTC 89
DB 221 CTTGACCCCTGGAGGGCTGCTCCATCGTGACACCACTTCCGCGTGGACCATCTT 280

QY 90 CCGGGTGTTACCATGTTCACGAAGCTGCTCCCACTCAAGCATAGTGTATGAGGACGGGA 149
DB 281 CGCGCTGTGGCCGGCGGACAAAGTGACCAACATCACCGCGCGGAGCAGTGGCTGCCGG 340

```

```

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGGGCAACTCCTCCCGTTG 209
DB 341 CATGGCATCTAGGTCCCGCACCGCTGTTCTGTCATTGCCCTGAAGGACGCCCCGGCTG 400

QY 210 C 210
DB 401 C 401

RESULT 10
AV637328
LOCUS
DEFINITION AV637328 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
CDNA clone HC071f12_r 5', mRNA sequence.
ACCESSION AV637328
VERSION
KEYWORDS
SOURCE
ORGANISM Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 451)
AUTHORS Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL DNA Res. 7 (5), 305-307 (2000)
MEDLINE 20539644
PUBMED 11089912
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
1..451
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC071f12_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT 74 a 150 c 150 g 77 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 451;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTGCTCTGCTGACCATACAGTTCGCTTATGAAGTGGCAACGTGTC 89
DB 178 CTTGACCCCTGGAGGGCTGCTCCATCGTGACACCACTTCCGCGTGGGACCATCTT 237

QY 90 CCGGGTGTTACCATGTTCACGAAGCTGCTCCCACTCAAGCATAGTGTATGAGGACGGGA 149
DB 238 CGCGCTGTGGCCGGCGGACAAAGTGACCAACATCACCGCGCGGAGCAGTGGCTGCCGG 297

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCGTTCGGGAGGGCAACTCCTCCCGTTG 209
DB 298 CATGGCATCTAGGTCCCGCACCGCTGTTCTGTCATTGCCCTGAAGGACGCCCCGGCTG 357

QY 210 C 210
DB 358 C 358

RESULT 11
AV637643
LOCUS

```



```
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC068a04_r"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      82 a 153 c 144 g 75 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 454;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTCTCTGCTGACCATACAGCTTCGGCTTATGAAGTGGCAACGTGTC 89
DB 52 CTTGACCCCTGGAGCGGCTGCTCCATCGTCGACACCAACTTCGCCGTGGGCACCATCTT 111

QY 90 CGGGGTGTACCATGTCCAGAACGACTGTCTCCAACTCAAGCATAGTGTATGAGGCAGCGGA 149
DB 112 CGCGGTGTGGCCGGCGGACAACTGACCAATCACCAGCGCGGAGCAGGTGGCTGCCGG 171

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCTCGGTTCGGGAGGGCAACTCTCTCCGTTG 209
DB 172 CATGGGCATCTACGGTCCCGCAGCCGCTGTTCTGCAATTGCGCTGAAGGAGCGCCCGCGCTG 231

QY 210 C 210
DB 232 C 232

RESULT 14
AV635382
LOCUS
DEFINITION
AV635382 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HC045f10_r 5', mRNA sequence.
ACCESSION
AV635382
VERSION
AV635382.1 GI:10778702
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Asamizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL
DNA Res. 7 (5), 305-307 (2000)
MEDLINE
20539644
PubMed
11089912
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Location/Qualifiers
source
1..456
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC045f10_r"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      75 a 158 c 148 g 75 t
ORIGIN

/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC068a04_r"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      82 a 153 c 144 g 75 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 454;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTCTCTGCTGACCATACAGCTTCGGCTTATGAAGTGGCAACGTGTC 89
DB 52 CTTGACCCCTGGAGCGGCTGCTCCATCGTCGACACCAACTTCGCCGTGGGCACCATCTT 111

QY 90 CGGGGTGTACCATGTCCAGAACGACTGTCTCCAACTCAAGCATAGTGTATGAGGCAGCGGA 149
DB 281 CGCGGTGTGGCCGGCGGACAACTGACCAATCACCAGCGCGGAGCAGGTGGCTGCCGG 340

QY 150 CATGATCATGCACACCCCGGGTGGTCCCTGCTCGGTTCGGGAGGGCAACTCTCTCCGTTG 209
DB 341 CATGGGCATCTACGGTCCCGCAGCCGCTGTTCTGCAATTGCGCTGAAGGAGCGCCCGCGCTG 400

QY 210 C 210
DB 401 C 401

RESULT 15
AV632765
LOCUS
DEFINITION
AV632765 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
cDNA clone HC012c10_r 5', mRNA sequence.
ACCESSION
AV632765
VERSION
AV632765.1 GI:10776085
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
1 (bases 1 to 473)
AUTHORS
Asamizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
TITLE
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
JOURNAL
DNA Res. 7 (5), 305-307 (2000)
MEDLINE
20539644
PubMed
11089912
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Location/Qualifiers
source
1..473
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="HC012c10_r"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      82 a 160 c 151 g 80 t
ORIGIN

Query Match      8.4%; Score 40.2; DB 9; Length 473;
Best Local Similarity 51.4%; Pred. No. 2.3;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 30 CTTGGCCCTGCTCTCTGCTGACCATACAGCTTCGGCTTATGAAGTGGCAACGTGTC 89
DB 156 CTTGACCCCTGGAGCGGCTGCTCCATCGTCGACACCAACTTCGCCGTGGGCACCATCTT 215

QY 90 CGGGGTGTACCATGTCCAGAACGACTGTCTCCAACTCAAGCATAGTGTATGAGGCAGCGGA 149
DB 216 CGCGGTGTGGCCGGCGGACAACTGACCAATCACCAGCGCGGAGCAGGTGGCTGCCGG 275
```

us-09-899-303a-11.rst

Mon Dec 22 13:28:34 2003

Qy	150	CATGATCATGCACACCCCGGGTGGTGCCTTCGGTTCGGGAGGCAACTCCTCCCGTTG	209
Db	276	CATGGGCATCTACGGTCCCGCACCGTGTCTGCATTGCCCTGAAGGACGCCCCCGGCTG	335
Qy	210	C	210
Db	336	C	336

Search completed: December 20, 2003, 06:54:43
Job time : 1234.83 secs

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	480	100.0	480	3	US-08-612-973-11	Sequence 11, Appl
2	480	100.0	480	3	US-08-927-597-11	Sequence 11, Appl
3	474.2	98.8	483	3	US-08-612-973-9	Sequence 9, Appl
4	474.2	98.8	483	3	US-08-927-597-9	Sequence 9, Appl
5	431.2	89.8	1539	2	US-08-470-426B-17	Sequence 17, Appl
6	431.2	89.8	1863	2	US-08-470-426B-14	Sequence 14, Appl
7	428	89.2	9595	3	US-09-014-416-4	Sequence 4, Appl
8	428	89.2	9599	3	US-09-014-416-6	Sequence 6, Appl
9	426.2	88.8	633	3	US-08-612-973-7	Sequence 7, Appl
10	426.2	88.8	633	3	US-08-927-597-7	Sequence 7, Appl
11	423.4	88.2	501	2	US-08-483-695-30	Sequence 30, Appl
12	423.4	88.2	501	2	US-07-965-285-30	Sequence 30, Appl
13	423.4	88.2	501	2	US-08-487-231-30	Sequence 30, Appl
14	423.4	88.2	501	3	US-09-201-913-30	Sequence 30, Appl
15	422.8	88.1	9472	4	US-08-150-204E-96	Sequence 96, Appl
16	422.2	88.0	642	3	US-08-612-973-3	Sequence 3, Appl
17	422.2	88.0	642	3	US-08-927-597-3	Sequence 3, Appl
18	421.2	87.8	795	3	US-08-612-973-5	Sequence 5, Appl
19	421.2	87.8	795	3	US-08-927-597-5	Sequence 5, Appl
20	421.2	87.8	2082	3	US-08-612-973-47	Sequence 47, Appl
21	421.2	87.8	2082	3	US-08-927-597-47	Sequence 47, Appl
22	421.2	87.8	2433	3	US-08-612-973-49	Sequence 49, Appl
23	421.2	87.8	2433	3	US-08-927-597-49	Sequence 49, Appl
24	419.6	87.4	1037	1	US-08-462-195-1	Sequence 1, Appl
25	419.6	87.4	1037	2	US-08-636-883-1	Sequence 1, Appl
26	419.6	87.4	1037	3	US-09-127-829-1	Sequence 1, Appl
27	417.2	86.9	1636	3	US-08-612-973-13	Sequence 13, Appl

```

1-08-612-973-11
Query Match 100.0%; Score 480; DB 3; Length 480;
Best Local Similarity 100.0%; Pred. No. 7,1e-126;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTCCGGTTGCTCTTTCTATCTTCTCTTGGCCCTGCTGCTCTGTCGACCATACCA 60
1 ATGTCCGGTTGCTCTTTCTATCTTCTCTTGGCCCTGCTGCTCTGTCGACCATACCA 60
61 GCTTCCGCTTATCAAGTGGCGCAACGTGTCCGGGTGTACCATGTACAGCAACGACTGCTCC 120
61 GCTTCCGCTTATGAAGTGGCGCAACGTGTCCGGGTGTACCATGTACAGCAACGACTGCTCC 120
121 AACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGTACACCCCGGGTGGTGCC 180
121 AACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGTACACCCCGGGTGGTGCC 180
181 TCGGTTCCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
181 TCGGTTCCGGAGGGCAACTCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
241 AGGAAGCGCAGCGTCCGCCAACACAAATACGACGCCACGTCGATTTGCTGTGGGGCT 300
241 AGGAAGCGCAGCGTCCGCCAACACAAATACGACGCCACGTCGATTTGCTGTGGGGCT 300
301 GCTGCTTTCTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTTCTTTGTTTCC 360
301 GCTGCTTTCTTCCGCTATGTACGTGGGGATCTCTCGGGATCTGTTTTCTTTGTTTCC 360
361 CAGCTGTTTCACTTTCTCACTCCCGGCATCAAAAGTACAGGACTGCAACTGCTCAATC 420
361 CAGCTGTTTCACTTTCTCACTCCCGGCATCAAAAGTACAGGACTGCAACTGCTCAATC 420
421 TATCCCGGCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGAACTGGTAATAG 480
421 TATCCCGGCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGAACTGGTAATAG 480

RESULT 2
JS-08-927-597-11
Sequence 11, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOF, GUY
APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.

```


Mon Dec 22 13:28:32 2003

```

Db 181 TCGGTTGGAGGGCAACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGC 240
Qy 241 AGGAAGCGCAGCGTCCCAACAACGACAAATACAGCCGACGTCGATTGCTGTTGGGGCT 300
Db 241 AGGAAGCGCAGCGTCCCAACAACGACAAATACAGCCGACGTCGATTGCTGTTGGGGCT 300
Qy 301 GCTGCTTCTGTTCCGCTATGTAAGTGGGGGATCTCTCGGATCTGTTCTTCTGTTTCC 360
Db 301 GCTGCTTCTGTTCCGCTATGTAAGTGGGGGATCTCTCGGATCTGTTCTTCTGTTTCC 360
Qy 361 CAGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGATGCAATGCTCAATC 420
Db 361 CAGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGATGCAATGCTCAATC 420
Qy 421 TATCCCGGCATGATATCAGTCAACCGATGCTTGGGATATGATGAACTGGTAATA 479
Db 421 TATCCCGGCATGATATCAGTCAACCGATGCTTGGGATATGATGAACTGGTCTTA 479

RESULT 5
US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-470-426B-17

Query Match 89.8%; Score 431.2; DB 2; Length 1539;
Best Local Similarity 94.1%; Pred. No. 5.1e-112;
Matches 448; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 2 TGTCCGGTTCCTTTCTTCTATCTCTGCTTGGCCCTGCTGCTGACATACACAG 61

```

```

Db 506 TGCCCGGTTGCTTTTCTTCTATCTTCTTCTGCTTGTCTGCTTTGACATCCCGAG 565
Qy 62 CTTCCGCTTATGAAGTGCACAGCTGTCCGGGGTGTACCATGTACAGAACGACTGCTCCA 121
Db 566 CTTCCGCTTATGAAGTGCACAGCTGTCCGGGGTGTACCATGTACAGAACGACTGCTCCA 625
Qy 122 ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGACATCTCCCGGGTGTGCTCCCT 181
Db 626 ACTCAAGCATAGTGTATGAGGAGCGGACATGATCATGACATCTCCCGGGTGTGCTCCCT 685
Qy 182 GCGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGCTCACTCCACGCTCGCGGCCA 241
Db 686 GCGTTCCGGAGGACAAACAGCTCCCGTTGCTGGGTAGGGTCACTCCACGCTCGCGGCCA 745
Qy 242 GGAACGCGAGGCTCCCAACAACGACAAATACGACGCGCATCAACAGTACAGGACTGCAATCT 301
Db 746 GGAATGCGAGGCTCCCAACAACGACAAATACGACGCGCATCAACAGTACAGGACTGCAATCT 805
Qy 302 CTGCTTCTGTTCCGCTATGTAAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
Db 806 CTGCTTCTGTTCCGCTATGTAAGTGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 865
Qy 362 AGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAATCTGCTCAATCT 421
Db 866 AGCTGTTCACTTCTCACTCGCGGCATCAACAGTACAGGACTGCAATCTGCTCAATCT 925
Qy 422 ATCCCGGCATGATATCAGTCAACCGATGCTTGGGATATGATGAACTGGTAA 477
Db 926 ATCCCGGCATGATATCAGTCAACCGATGCTTGGGATATGATGAACTGGTCA 981

RESULT 6
US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1863 base pairs
; TYPE: nucleic acid

```



```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..627
; US-08-927-597-7

Query Match      88.8%; Score 426.2; DB 3; Length 633;
Best Local Similarity 93.1%; Pred. No. 1e-110;
Matches 446; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2 TGTCCGGTGTCTTTCTCTATCTTCTCTTGTGGCCCTGCTGTCTGTGACCATACCAG 61
Db 155 TGCCTGTGTCTTTCTCTATCTTCTCTTGTGGCTTTACTGTCTGTGACCATCCAG 214
QY 62 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGGTACCATGTGCAGAACGACTGTCCA 121
Db 215 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGGTACCATGTGCAGAACGACTGTCCA 274
QY 122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 181
Db 275 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 334
QY 182 GGTTCGGAGGGCAACTCTCTCCGTTGTGGTGGCGTCACTCCACGCTCGCGGCA 241
Db 335 GCGTTCGGAGAACAACTCTTCCGCTGTGGGTAGCGCTCACCCCGCTCGCAGTA 394
QY 242 GGAACGCCAGCGTCCCGCAACAGACATACGACGACGACGTCGATTTGCTGGGGCTG 301
Db 395 GGAACGCCAGCGTCCCGCAACAGACATACGACGACGACGTCGATTTGCTGGGGCTG 454
QY 302 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTTTCC 361
Db 455 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTTTCC 514
QY 362 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db 515 AGCTGTTCACCTTCTCACCTCGCGGCATGAGAGCGGTGCAGGACTGCAATTGCTCAATCT 574
QY 422 ATCCCGGCATGTATCAGTCCAGTCACCGCATGGCTTGGGATATGATGAACTGGTAATAG 480
Db 575 ATCCCGGCATGTATCAGTCCAGTCCAGTCACCGCATGGCTTGGGATATGATGAACTGGTAATAG 633

RESULT 11
US-08-483-695-30
; Sequence 30, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 08-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; DESCRIPTION: cDNA to genomic RNA
; US-08-483-695-30

Query Match      88.2%; Score 423.4; DB 2; Length 501;
Best Local Similarity 94.4%; Pred. No. 5.8e-110;
Matches 439; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 2 TGTCCGGTGTCTTTCTCTATCTTCTCTTGTGGCCCTGCTGTCTGTGACCATACCAG 61
Db 37 TGCCTGTGTCTTTCTCTATCTTCTCTTGTGGCTTTGCTGTCTGTGACCATCCAG 96
QY 62 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGGTGACCATGTGCAGAACGACTGTCCA 121
Db 97 CTTCCGCTTATGAAGTGGCGCAACGTGTCGGGGGTGACCATGTGCAGAACGACTGTCCA 156
QY 122 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTCCCT 181
Db 157 ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCATACTCCCGGGTGGTCCCT 216
QY 182 GCGTTCGGAGGGCAACTCTCTCCGTTGTGGTGGCGTCACTCCACGCTCGCGGCA 241
Db 217 GCGTTCGGAGGGCAACAGCTCCCGTTGCTGGGTAGCGTCACTCCACGCTCGCGGCA 276
QY 242 GGAACGCCAGCGTCCCGCAACAGACAAATAGACGCCACGTCGATTTGCTGGGGCTG 301
Db 277 GGAATGCCAGCGTCCCGCAACAGACAAATAGACGCCACGTCGATTTGCTGGGGCTG 336
QY 302 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTTTCC 361
Db 337 CTGCTTTCTGTTCCGCTATGTAGTGGGGATCTCTGCGGATCTGTTTCTTTTCC 396
QY 362 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db 397 AGCTGTTCACCTTCTCACCTCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 456
QY 422 ATCCCGGCATGTATCAGTCCAGTCCAGTCACCGCATGGCTTGGGATATGATGA 466
Db 457 ATCCCGGCATGTATCAGTCCAGTCCAGTCACCGCATGGCTTGGGATATGATGA 501

RESULT 12
US-07-965-285-30
; Sequence 30, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
```


us-09-899-303a-11.rni

Mon Dec 22 13:28:32 2003

```

APPLICATION NUMBER: KR 91-13601
FILING DATE: 6-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Shahan Ielam, Esq.
REGISTRATION NUMBER: 32,507
REFERENCE/DOCKET NUMBER: 2695/FLK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 940-8564
TELEFAX: (212) 940-8776
INFORMATION FOR SEQ ID NO: 96
SEQUENCE CHARACTERISTICS:
LENGTH: 9472 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: KHCV-LBC1, Fig. 2
SEQUENCE DESCRIPTION: SEQ ID NO: 96
US-08-150-204E-96

Query Match      88.1%; Score 422.8; DB 4; Length 9472;
Best Local Similarity 93.2%; Pred. No. 2e-109;
Matches 442; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY      2  TGTCCGGTTGCTCTTCTCTATCTTCTCTTGGCCCTGTCTGTCTGTGACCATACACAG 61
Db      848  TGTCCGGTTGCTCTTCTCTATCTTCTCTTGGCTGTCTGTCTGTGACCATACACAG 907

QY      62  CTTCCGGTTATGAAGTCGCAACGTTGCCGGGTGTACCATGTACGACGACGACTGCTCCA 121
Db      908  TTTCCGGTTATGAAGTCGCTAAGCGTCCGGGATGTACCATGTACGACGACGACTGCTCCA 967

QY      122  ACTCAAGCATAGTGTATGAGGACGCGACATGATCATGCACACCCCGGGTGGTGCCT 181
Db      968  ACTCAAGCATTTGTGTATGAGGACGCGACATGATCATGCACACTCCCGGGTGGTGCCT 1027

QY      182  GCGTTCCGGGAGGCAACTCTCTCCCGTTGCTGGGTGGCGCTCATCTCCACGCTCGCGCCA 241
Db      1028  GCGTTCCGGGAGGACAACTCTCTCCCGTTGCTGGGTGGCACTTACTCCACGCTCGCGCCA 1087

QY      242  GGAACGCCAGCGTCCCAACAGACATACAGCCGACGTCGATTGCTCGTTGGGGCTG 301
Db      1088  GGAATGCCAGCGTCCCACTAGACATTGCCAGCCGATGTCGATTGCTCGTTGGGGTAG 1147

QY      302  CTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGATCTGTTTCCTTTTCCC 361
Db      1148  CTGCTTTCTGTTCCGCTATGTACGTGGGGGACCTCTGCGATCTGTTTCCTTTTCCC 1207

QY      362  AGCTGTTCACCTTCTCACTCGCCGGGATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db      1208  AGCTGTTCACCTTTTGGCTTCGCCGATGAGACGCTACAGGACTGCAACTGCTCAATCT 1267

QY      422  ATCCCGGCCATGTATCAGTACCGCATCGGCTTGGGATATGATGATGAATGGT 475
Db      1268  ATCCCGGCCGCGTATCAGGTACCGCATGGCTTGGGATATGATGATGAATGGT 1321
```

Search completed: December 20, 2003, 07:03:01
Job time : 35.6174 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2408.07 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-13
Perfect score: 636
Sequence: 1 ATGCTGGTGAAGCCATCGA.....TGATGACTGGTACTAATAAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues 5777422
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hg_hum.*
- 31: em_hg_inv.*
- 32: em_hg_other.*
- 33: em_hg_mus.*
- 34: em_hg_pln.*
- 35: em_hg_rod.*
- 36: em_hg_mam.*
- 37: em_hg_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	636	100.0	636	6	A48675	A48675 Sequence 13
2	636	100.0	636	6	AR157329	AR157329 Sequence
3	636	100.0	636	6	AX452762	AX452762 Sequence
4	636	100.0	636	6	AX685014	AX685014 Sequence
5	612.8	96.4	633	6	A48669	A48669 Sequence 7
6	612.8	96.4	633	6	AR157326	AR157326 Sequence
7	612.8	96.4	633	6	AX452756	AX452756 Sequence
8	612.8	96.4	633	6	AX685008	AX685008 Sequence
9	604	95.0	795	6	A48667	A48667 Sequence 5
10	604	95.0	795	6	AR157325	AR157325 Sequence
11	604	95.0	795	6	AX452754	AX452754 Sequence
12	604	95.0	795	6	AX685006	AX685006 Sequence
13	601.6	94.6	2082	6	A48709	A48709 Sequence 47
14	601.6	94.6	2082	6	AR157350	AR157350 Sequence
15	601.6	94.6	2082	6	AX452796	AX452796 Sequence
16	601.6	94.6	2082	6	AX685048	AX685048 Sequence
17	601.6	94.6	2433	6	A48711	A48711 Sequence 49
18	601.6	94.6	2433	6	AR157351	AR157351 Sequence
19	601.6	94.6	2433	6	AX452798	AX452798 Sequence
20	601.6	94.6	2433	6	AX685050	AX685050 Sequence
21	566.4	89.1	9379	14	AF207766	AF207766 Hepatitis
22	564.8	88.8	9386	14	AF165056	AF165056 Hepatitis
23	563.2	88.6	9379	14	HCVPOLYP	AJ000009 Hepatitis
24	563.2	88.6	9386	14	AF165055	AF165055 Hepatitis
25	560.6	88.1	9359	14	AF313916	AF313916 Hepatitis
26	560	88.1	9379	14	AF165047	AF165047 Hepatitis
27	560	88.1	9379	14	AF165051	AF165051 Hepatitis
28	560	88.1	9379	14	AF165052	AF165052 Hepatitis
29	560	88.1	9598	14	AB049101	AB049101 Hepatitis
30	558.4	87.8	9379	14	AF165048	AF165048 Hepatitis
31	558.4	87.8	9410	14	HPCCKR2	D50481 Hepatitis C
32	556.8	87.5	1880	14	HPCSTRJ4	D00832 Hepatitis C
33	556.8	87.5	2540	6	E04260	E04260 cDNA encodi
34	556.8	87.5	2540	6	E04805	E04805 cDNA to 5'
35	556.8	87.5	2540	6	E07391	E07391 cDNA encodi
36	556.8	87.5	9344	14	AB049096	AB049096 Hepatitis
37	556.8	87.5	9379	14	AF207761	AF207761 Hepatitis
38	556.8	87.5	9448	14	HPCJ483	D13558 Hepatitis C
39	556.8	87.5	9605	6	AX739971	AX739971 Sequence
40	556.8	87.5	9605	14	HCJ238799	AJ238799 Hepatitis
41	556.8	87.5	11076	6	AX036252	AX036252 Sequence
42	556.8	87.5	11076	6	AX036258	AX036258 Sequence
43	556.8	87.5	11076	6	AX036260	AX036260 Sequence
44	556.8	87.5	11076	6	AX036262	AX036262 Sequence
45	556.2	87.5	9431	14	HCJ45476	U45476 Hepatitis C

ALIGNMENTS

RESULT 1
A48675
LOCUS
DEFINITION Sequence 13 from Patent WO9604385.
ACCESSION A48675
VERSION A48675.1 GI:2302388
KEYWORDS
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 636)
Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 13 15-FEB-1996;

Mon Dec 22 13:28:34 2003

INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.

FEATURES

source
1..636
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..633
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03133.1"
/db_xref="GI:2302389"
/translation="MLGKAI DLTTCGFADLVGYIPLVGA PLGGARALAHGVRVLEQ
VNYATNLPGCSFISIFLLALLSCLIPASAYEVNVSGMVHTNDCNSSIIVGEADM
IHMTGCVPCVRENNSSRCWVALTPGTCVRENNSSRIPTTIRRHVDLVGAAAFCSAMVY
GDLGCVFLVLSQLTFTISPRHETVQDNCNSIYPGHTTGHMAWMMWNY"

CDS

1..630
/product="unnamed"
mat_peptide 111 a 197 c 175 g 153 t

BASE COUNT 111 a 197 c 175 g 153 t

ORIGIN

Query Match 100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 5.5e-137;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGGGTAAGGCATCGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGCTGGGTAAGGCATCGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
DB 61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
QY 121 CTGGAAGACGGCGTGAACATATGCAACAGGGAAATTTGGCTGGTCTCTTCTATCTTC 180
DB 121 CTGGAAGACGGCGTGAACATATGCAACAGGGAAATTTGGCTGGTCTCTTCTATCTTC 180
QY 181 CTCCTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCGCTTCTTCTATCTTC 240
DB 181 CTCCTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCGCTTCTTCTATCTTC 240
QY 241 TCCGGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCAACACCCCGGGTGGTGGCGGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCAACACCCCGGGTGGTGGCGGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGTCAACCCCGGCTAGGAAAGCGGCTAGGAAAGCGGATCCCACTACACA 420
DB 361 TCGTGGGTAGCGTCAACCCCGGCTAGGAAAGCGGCTAGGAAAGCGGATCCCACTACACA 420
QY 421 ATACGACGCGCATCGATTGGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
DB 421 ATACGACGCGCATCGATTGGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGCGGATCTGTCTTCTCTGCTCCAGCTGTTCCAGCTTCGCGGTCAACGT 540
DB 481 GGGGATCTCTGCGGATCTGTCTTCTCTGCTCCAGCTGTTCCAGCTTCGCGGTCAACGT 540
QY 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGCACATAACGGGTCAACCGT 600
DB 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGCACATAACGGGTCAACCGT 600
QY 601 ATGGCTTGGGATATGATGATGAATGTAATAG 636
DB 601 ATGGCTTGGGATATGATGATGAATGTAATAG 636

RESULT 2
AR157329

LOCUS AR157329 636 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 13 from patent US 6245503.
ACCESSION AR157329
VERSION AR157329.1 GI:16218262
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 636)
AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: US 6245503-A 13 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..636
BASE COUNT 111 a 197 c 175 g 153 t
ORIGIN

Query Match 100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 5.5e-137;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGGGTAAGGCATCGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGCTGGGTAAGGCATCGATACCCCTTACGTCGGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
DB 61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT 120
QY 121 CTGGAAGACGGCGTGAACATATGCAACAGGGAAATTTGGCTGGTCTCTTCTATCTTC 180
DB 121 CTGGAAGACGGCGTGAACATATGCAACAGGGAAATTTGGCTGGTCTCTTCTATCTTC 180
QY 181 CTCCTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCGCTTCTTCTATCTTC 240
DB 181 CTCCTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCGCTTCTTCTATCTTC 240
QY 241 TCCGGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCAACACCCCGGGTGGTGGCGGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCAACACCCCGGGTGGTGGCGGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGTCAACCCCGGCTAGGAAAGCGGCTAGGAAAGCGGATCCCACTACACA 420
DB 361 TCGTGGGTAGCGTCAACCCCGGCTAGGAAAGCGGCTAGGAAAGCGGATCCCACTACACA 420
QY 421 ATACGACGCGCATCGATTGGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
DB 421 ATACGACGCGCATCGATTGGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGCGGATCTGTCTTCTCTGCTCCAGCTGTTCCAGCTTCGCGGTCAACGT 540
DB 481 GGGGATCTCTGCGGATCTGTCTTCTCTGCTCCAGCTGTTCCAGCTTCGCGGTCAACGT 540
QY 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGCACATAACGGGTCAACCGT 600
DB 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGCACATAACGGGTCAACCGT 600
QY 601 ATGGCTTGGGATATGATGATGAATGTAATAG 636
DB 601 ATGGCTTGGGATATGATGATGAATGTAATAG 636

RESULT 3
AX452762 636 bp DNA linear PAT 06-JUL-2002
LOCUS AX452762
DEFINITION Sequence 13 from Patent EP1211315.
ACCESSION AX452762

```
VERSION      AX452762.1  GI:21712447
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Hepacivirus.

REFERENCE    1
AUTHORS      Maertens, G., Bosman, F., de Martynoff, G. and Buysse, M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 121315-A 13 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..636
                /organism="Hepatitis C virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:11103"
                1..633
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD38071.1"
                /db_xref="GI:21712448"
                /translation="MLGKAIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRVLEDG
                VNYATGNLPCVCSFSLFLALLSLTIPASAYERNVSGMYHTNDSCNSSIIVVEADM
                IMHTPCVPCVENNSSRCWALPTPLAARNASIPPTTIRRHVDLLVGAACFCSAMYV
                GDLCSVFLVSQLFTISPRHETVQDCNCSIYPGHITGHRMAMMMWNY"
              mat_peptide
                1..630
                /product="unnamed"
              BASE COUNT  111 a  197 c  175 g  153 t
              ORIGIN
                100.0%; Score 636; DB 6; Length 636;
                Best Local Similarity 100.0%; Pred. No. 5.5e-137;
                Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGCGTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGCGTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGGCCCTTAGGGGCGCTGCCAGGCGCTGCCGATGGCGTTCGGGTT 120
DB 61 CCGCTCGTGGCGGCCCTTAGGGGCGCTGCCAGGCGCTGCCGATGGCGTTCGGGTT 120

QY 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCGCTGCTTCTCTATCTTC 180
DB 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCGCTGCTTCTCTATCTTC 180

QY 181 CTCCTGGCTTTACTGCTGCTTAAACCATTCAGCTTCCGCTTACGAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTACTGCTGCTTAAACCATTCAGCTTCCGCTTACGAGTGGCAACGTG 240

QY 241 TCCGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTCGTATGAGGAGCG 300
DB 241 TCCGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTCGTATGAGGAGCG 300

QY 301 GACATGATCATGCACACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 360

QY 361 TCGTGGGTAGCGTCAACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 420
DB 361 TCGTGGGTAGCGTCAACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 420
```

```
QY 601 ATGGCTTGGATATGATGATGAAGTGGTACTAATAG 636
DB 601 ATGGCTTGGATATGATGATGAAGTGGTACTAATAG 636

RESULT 4
AX452762.1  GI:21712447
LOCUS       AX452762.1  636 bp  DNA  linear  PAT 29-MAR-2003
DEFINITION  Sequence 13 from Patent WO0205548.
ACCESSION  AX452762.1
VERSION    AX452762.1  GI:21712447
KEYWORDS   Hepatitis C virus
SOURCE     Hepatitis C virus
           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
           Hepacivirus.
REFERENCE  1
AUTHORS    Maertens, G., Bosman, F. and Buysse, M.A.
TITLE      Purified Hepatitis C Virus envelope proteins for diagnostic and
           therapeutic use
JOURNAL    Patent: WO 0205548-A 13 18-JUL-2002;
           INNOGENETICS N.V. (BE)
FEATURES   Location/Qualifiers
           source
             1..636
             /organism="Hepatitis C virus"
             /mol_type="genomic DNA"
             /db_xref="taxon:11103"
             1..633
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="CAD86524.1"
             /db_xref="GI:29371420"
             /translation="MLGKAIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRVLEDG
             VNYATGNLPCVCSFSLFLALLSLTIPASAYERNVSGMYHTNDSCNSSIIVVEADM
             IMHTPCVPCVENNSSRCWALPTPLAARNASIPPTTIRRHVDLLVGAACFCSAMYV
             GDLCSVFLVSQLFTISPRHETVQDCNCSIYPGHITGHRMAMMMWNY"
           mat_peptide
             1..630
             /product="unnamed"
           BASE COUNT  111 a  197 c  175 g  153 t
           ORIGIN
```

```
Query Match 100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 5.5e-137;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGCGTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGGCGTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGGCCCTTAGGGGCGCTGCCAGGCGCTGCCGATGGCGTTCGGGTT 120
DB 61 CCGCTCGTGGCGGCCCTTAGGGGCGCTGCCAGGCGCTGCCGATGGCGTTCGGGTT 120

QY 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCGCTGCTTCTCTATCTTC 180
DB 121 CTGGAAGACGGGTGAACATGCAACAGGGAATTCGCTGCTTCTCTATCTTC 180

QY 181 CTCCTGGCTTTACTGCTGCTTAAACCATTCAGCTTCCGCTTACGAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTACTGCTGCTTAAACCATTCAGCTTCCGCTTACGAGTGGCAACGTG 240

QY 241 TCCGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTCGTATGAGGAGCG 300
DB 241 TCCGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTCGTATGAGGAGCG 300

QY 301 GACATGATCATGCACACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 360

QY 361 TCGTGGGTAGCGTCAACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 420
DB 361 TCGTGGGTAGCGTCAACCCCGGTCGTCCTCGCTTCCGAGAACAACTCTTCCCGC 420
```

QY	421	ATACAGCGCAGCTCGATTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG	480
Db	421	ATACAGCGCAGCTCGATTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG	480
QY	481	GGGGATCTCTGGGATCTGTCTCTCGTCTCCAGCTGTTCCACATCTCGCTCGCGG	540
Db	481	GGGGATCTCTGGGATCTGTCTCTCGTCTCCAGCTGTTCCACATCTCGCTCGCGG	540
QY	541	CATGAGACGGTGCAGAGCTGCAATGTCTCAATCTATCCCGGCCACATACGGGTCA	600
Db	541	CATGAGACGGTGCAGAGCTGCAATGTCTCAATCTATCCCGGCCACATACGGGTCA	600
QY	601	ATGGCTTGGGATATGATGATGAATGGTACTAATAG	636
Db	601	ATGGCTTGGGATATGATGATGAATGGTACTAATAG	636
RESULT 5			
LOCUS	A48669	633 bp	DNA linear PAT 07-MAR-1997
DEFINITION	Sequence 7 from Patent WO9604385.		
ACCESSION	A48669		
VERSION	A48669.1	GI:2302382	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 633)		
AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buysee,M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	Patent: WO 9604385-A 7 15-FEB-1996;		
COMMENT	INNOGENETICS NV (BE)		
	Other publication CA 2172273 960215		
	Other publication AU 338495 960304.		
FEATURES	Location/Qualifiers		
source	1..633		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
CDS	1..630		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAA03130.1"		
	/db_xref="GI:2302383"		
	/translation="MLGKVDTLTCGFADIMGYIPLVGAPLGAARALAHGVRVLEDG VNYATNLPGCSFSIFLALLSLTTPAGAYEVNRVNGMYHTNDCSNSSIVYEAADM IMHTPGCVPCVRENNSSECHVALTPTLAARNASVPPTTIRRHVDLLVGAAPFCAMYY GDLCGSFVLSQLFTIISPRRHETVQDCNCSLIYPGHITGRMAWDMWNN"		
mat_peptide	1..627		
	/product="unnamed"		
BASE COUNT	111 a 192 c 174 g 156 t		
ORIGIN			
Query Match	96.4%;	Score 612.8;	DB 6; Length 633;
Best Local Similarity	98.1%;	Pred. No. 1.3e-131;	
Matches	620; Conservative	0; Mismatches	12; Indels 0; Gaps 0;
QY	1	ATGCTGGGTAAAGCCATCGATACCCCTTACGTGCGGCTTTCGCGACCTCTGTGGGGTACATT	60
Db	1	ATGTTGGGTAAAGTCAATCGATACCCCTTACGTGCGGCTTTCGCGACCTCTGTGGGGTACATT	60
QY	61	CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCAGAGCCCTGGCGCATGCGCTCCGGGTT	120
Db	61	CCGCTCGTGGGGCCCCCTAGGGGGGTGCTGCAGAGCCCTGGCGCATGCGCTCCGGGTT	120
QY	121	CTGGAAGACGGCGTGAACCTATGCAACAGGAAATTTGCCCTGTTCTCTATCTTC	180
Db	121	CTGGAAGACGGCGTGAACCTATGCAACAGGAAATTTGCCCTGTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCCGCTTACGAGTGGCAACGTG	240
Db	181	CTCTTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCCGCTTATGAGTGGCAACGTG	240
QY	181	CTCTTGGCTTTACTGTCCTGTCTTAACCAATTCAGCTTCCGCTTACGAGTGGCAACGTG	300

```
Db 241 TCCTGGATGATACATGCTACGACGAGCTGCTCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGTCAACCCCGAGCTCGCGGTAGGAAGCCAGCATCCCACTACAACA 420
Db 361 TCGTGGGTAGCGTCAACCCCGAGCTCGCGGTAGGAAGCCAGCATCCCACTACAACA 420
QY 421 ATACGACGCCACGCTCGATTGCTGCTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGCTCGATTGCTGCTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGGGATCTGCTTCTCCAGCTGTGTTACCATCTCGCTCGCGCG 540
Db 481 GGGGATCTCTGGGATCTGCTTCTCCAGCTGTGTTACCATCTCGCTCGCGCG 540
QY 541 CATGACAGCGTGCAGACTGCAATTCCTCAATCTATCCCGGCCACATAAGCGGTACCGT 600
Db 541 CATGACAGCGTGCAGACTGCAATTCCTCAATCTATCCCGGCCACATAAGCGGTACCGT 600
QY 601 ATGGCTTGGGATATGATGATGAACCTGTAATA 632
Db 601 ATGGCTTGGGATATGATGATGAACCTGTAATA 632

RESULT 7
AX452756
LOCUS AX452756 633 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 7 from Patent EP1211315.
ACCESSION AX452756
VERSION AX452756.1 GI:21712441
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buysse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 7 05-JUN-2002;
INNOCENETICS N.V. (BE)
FEATURES
source
Location/Qualifiers
1..633
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:111103"
1..630
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD38068.1"
/db_xref="GI:21712442"
VNYATNLPKCSFIFLLALLSLTIPASAYEVNVSVMYHVTNDCSNSSIVYEAADM
INHTPCGVCVRNENSSRCWALTPTLAARNAASVPTTIRRHVDLLVGAAGFCSAMVY
GDLGSGVFLVSLFTISPRRHETVQDCNCSIPGHITGHRMAMWMMNWN"
mat_peptide 1..627
/product="unnamed"
BASE COUNT 111 a 192 c 174 g 156 t
ORIGIN
Query Match 96.4%; Score 612.8; DB 6; Length 633;
Best Local Similarity 98.1%; Pred. No. 1.3e-131;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATCTGGGTAAAGCCATCGATACCTTACGTGCGGCTTCGCCACCTCGTGGGTACATT 60
Db 1 ATCTGGGTAAAGTCATCGATACCTTACGTGCGGCTTCGCCACCTCATGGGTACATT 60
QY 61 CCCTCGTCCGCGCCCTTACGGGGCGCTGCCAGGCGCTTCCGCGCATGCGGTCCGGTT 120
```

```
Db 61 CCGCTCGTCCGCGCCCTTACGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGTT 120
QY 121 CTGGAAGACGGGTGAATCTATGCAACAGGGAATTTGCCTGGTTGCTCTTTCTATCTTC 180
Db 121 CTGGAAGACGGGTGAATCTATGCAACAGGGAATTTGCCTGGTTGCTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTACTGCTCTGCTTAACCATTCAGCTTCCGCTTACAGGTGGCAACGTG 240
Db 181 CTCTTGGCTTTACTGCTCTGCTTAACCATTCAGCTTCCGCTTATGAGGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTACGAACGACTGCTCCAATCAAGCATTTGTATGAGGAGCG 300
Db 241 TCCGGGATGTACCATGTACGAACGACTGCTCCAATCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGTCAACCCCGAGCTCGCGGTAGGAAGCCAGCATCCCACTACAACA 420
Db 361 TCGTGGGTAGCGTCAACCCCGAGCTCGCGGTAGGAAGCCAGCATCCCACTACAACA 420
QY 421 ATACGACGCCACGCTCGATTGCTGCTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGCTCGATTGCTGCTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGGGATCTGCTTCTCCAGCTGTGTTACCATCTCGCTCGCGCG 540
Db 481 GGGGATCTCTGGGATCTGCTTCTCCAGCTGTGTTACCATCTCGCTCGCGCG 540
QY 541 CATGACAGCGTGCAGACTGCAATTCCTCAATCTATCCCGGCCACATAAGCGGTACCGT 600
Db 541 CATGACAGCGTGCAGACTGCAATTCCTCAATCTATCCCGGCCACATAAGCGGTACCGT 600
QY 601 ATGGCTTGGGATATGATGATGAACCTGTAATA 632
Db 601 ATGGCTTGGGATATGATGATGAACCTGTAATA 632

RESULT 8
AX685008
LOCUS AX685008 633 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 7 from Patent WO0205548.
ACCESSION AX685008
VERSION AX685008.1 GI:29371413
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F. and Buysse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: WO 0205548-A 7 18-JUL-2002;
INNOCENETICS N.V. (BE)
FEATURES
source
Location/Qualifiers
1..633
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:111103"
1..630
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD86521.1"
/db_xref="GI:29371414"
VNYATNLPKCSFIFLLALLSLTIPASAYEVNVSVMYHVTNDCSNSSIVYEAADM
INHTPCGVCVRNENSSRCWALTPTLAARNAASVPTTIRRHVDLLVGAAGFCSAMVY
GDLGSGVFLVSLFTISPRRHETVQDCNCSIPGHITGHRMAMWMMNWN"
mat_peptide 1..627
/product="unnamed"
BASE COUNT 111 a 192 c 174 g 156 t
```


481 GGCGATCTTCGGGATCTGTCTTCTCTCCAGCTGTTCCACCATCTCGCTCGCCGG 540
481 GGCGACCTCTGGCGGATCTGTCTTCTCTCCAGCTGTTCCACCATCTCGCTCGCCGG 540
541 CATGAGACGGTGCAGGACTGCAATTGCTCAATTCATCCCGGCACATAACGGGTACCGT 600
541 CATGAGACGGTGCAGGACTGCAATTGCTCAATTCATCCCGGCACATAACGGGTACCGT 600
601 ATGGCTTGGGATATGATGATGAACCTGGT 628
601 ATGGCTTGGGATATGATGATGAACCTGGT 628

RESULT 13
A48709
LOCUS A48709 2082 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 47 from Patent WO9604385.
ACCESSION A48709
VERSION A48709.1 GI:2302422
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2082)
AUTHORS Maertens, G., Bosman, F., De, M. G. and Buyse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL INNOGENETICS NV (BE)
COMMENT Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
source
1..2082
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..2079
/note="unnamed protein product; Protein sequence is in conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA03146.1"
/translation="NLGKVIDTLTCTGFPADLVGYIPLVGLAPLGGARALAHGVRVLEDD
VNYATGNLPGCSFISFILLALLSCLTVPASAYEVNRVSGMVHTVNDSCNSIIVEAADM
IMHTPGCVPCVRENSSRCWALTPTLAARNASVPTTIRRHVDLLVGAAPFCSAMVY
GDLCSFVLVSQFTISPRRHETVQDCNCSIYPGHIITGHRMADMMWSPPTALVVS
QLLRIPQAVDMVAGAHGVLGLAYYSMVGNWAKVLIVMLLPAP"
mat_peptide 1..2076
/product="unnamed"
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN
Query Match 94.6%; Score 601.6; DB 6; Length 2082;
Best Local Similarity 97.8%; Pred. No. 5.3e-129;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
5 TGGGTAAAGGCATCATGATACCTTACGTCGGCTTCGCGACCTCGTGGGGTACATTCGCG 64
5 TGGGTAAAGGCATCATGATACCTTACGTCGGCTTCGCGACCTCGTGGGGTACATTCGCG 64
65 TCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGGTTCTGG 124
65 TCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGGTTCTGG 124

AX685006 795 bp DNA linear PAT 29-MAR-2003
LOCUS AX685006 795 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 5 from Patent WO02055548.
ACCESSION AX685006
VERSION AX685006.1 GI:29371411
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Hepatitis C virus
Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: WO 02055548-A 5 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES
source
1..795
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..792
/note="unnamed protein product"
/codon_start=1
/protein_id="CND86520.1"
/translation="NLGKVIDTLTCTGFPADLVGYIPLVGLAPLGGARALAHGVRVLEDD
VNYATGNLPGCSFISFILLALLSCLTVPASAYEVNRVSGMVHTVNDSCNSIIVEAADM
IMHTPGCVPCVRENSSRCWALTPTLAARNASVPTTIRRHVDLLVGAAPFCSAMVY
GDLCSFVLVSQFTISPRRHETVQDCNCSIYPGHIITGHRMADMMWSPPTALVVS
QLLRIPQAVDMVAGAHGVLGLAYYSMVGNWAKVLIVMLLPAP"
mat_peptide 1..789
/product="unnamed"
BASE COUNT 130 a 240 c 231 g 194 t
ORIGIN
Query Match 95.0%; Score 604; DB 6; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.4e-129;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 ATGTGGGTAAAGGCATCATGATACCTTACGTCGGCTTCGCGACCTCGTGGGGTACATT 60
1 ATGTGGGTAAAGGCATCATGATACCTTACGTCGGCTTCGCGACCTCGTGGGGTACATT 60
61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
121 CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTTC 180
121 CTGGAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTTC 180
181 CTCTGGCTTACTGTCCTGTCTAACAATTCAGCTTCCAGCTTACGAGTGGCAACGTG 240
181 CTCTGGCTTACTGTCCTGTCTAACAATTCAGCTTCCAGCTTACGAGTGGCAACGTG 240
241 TCCGGGATGTACCATGTACGACGACGATGCTCAACTAAGCATTTGTATGAGGACGG 300
241 TCCGGGATGTACCATGTACGACGACGATGCTCAACTAAGCATTTGTATGAGGACGG 300
301 GACATGATCATGACACCCCGGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
301 GACATGATCATGACACCCCGGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
361 TGCTGGGTAGCGGTCAACCCCGACGCTCGCGGTAGGAACGCGAGCATCCCACTCAACA 420
361 TGCTGGGTAGCGGTCAACCCCGACGCTCGCGGTAGGAACGCGAGCATCCCACTCAACA 420
421 ATACAGCGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTTCTCGGTATGTACGTG 480
421 ATACAGCGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTTCTCGGTATGTACGTG 480

Mon Dec 22 13:28:34 2003

```
mat_peptide 1..2076
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN
Query Match 94.6%; Score 601.6; DB 6; Length 2082;
Best Local Similarity 97.8%; Pred. No. 5.3e-129;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGGTAAGGCCATCGATACCTTACGTGGGCTTCGCCGACCTCGTGGGTACATTCCGC 64
Db |||||||
QY 5 TGGGTAAGGTCATCGATACCTTACATCGGGCTTCGCCGACCTCGTGGGTACATTCCGC 64
Db |||||||
QY 65 TCCTCGGCGGCCCTTAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTGG 124
Db |||||||
QY 65 TCCTCGGCGGCCCTTAGGGGGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTGG 124
Db |||||||
QY 125 AAGACGGCGTGAACATATGCAACAGGGAATTGGCTGGTTGCTCTTCTATCTCTCT 184
Db |||||||
QY 125 AAGACGGCGTGAACATATGCAACAGGGAATTGGCGGGTGGCTCTTCTATCTCTCT 184
Db |||||||
QY 185 TGGCTTTATCTGCTGTCTAACTTCCAGCTTCCGCTTACGAGTGGCAACGTGTCCG 244
Db |||||||
QY 185 TGGCTTTATCTGCTGTCTAACTTCCAGCTTCCGCTTACGAGTGGCAACGTGTCCG 244
Db |||||||
QY 245 GGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCAATTGTATGAGGAGCGGACA 304
Db |||||||
QY 245 GGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCAATTGTATGAGGAGCGGACA 304
Db |||||||
QY 305 TGATCATGCAACACCCCGGGTGGTCCCTGGCTTCGGGAGAACAACTCTTCCCGCTGCT 364
Db |||||||
QY 305 TGATCATGCAACACCCCGGGTGGTCCCTGGCTTCGGGAGAACAACTCTTCCCGCTGCT 364
Db |||||||
QY 365 GGGTAGCGCTCACCCCGGCTCGCGGTAGGAACGCGCATCCCCACCTACCAACAATAC 424
Db |||||||
QY 365 GGGTAGCGCTCACCCCGGCTCGCGGTAGGAACGCGCATCCCCACCTACCAACAATAC 424
Db |||||||
QY 425 GACGCACGTGCAITTTGCTCGTTGGGGCGGTGCTTTCTGTTCCGCTATGATGTTGGGG 484
Db |||||||
QY 425 GACGCACGTGCAITTTGCTCGTTGGGGCGGTGCTTTCTGTTCCGCTATGATGTTGGGG 484
Db |||||||
QY 485 ATCTCTGGGATCTGTCTTCTCTCTCCAGCTGTTCAACCATCTCGCTCCCGGCGATG 544
Db |||||||
QY 485 ACCTCTGGGATCTGTCTTCTCTCTCCAGCTGTTCAACCATCTCGCTCCCGGCGATG 544
Db |||||||
QY 545 AGACGGTGAGGACTGCAATTGCTCAATCTATCCCGGCGCAATAACGGGTACCGTATGG 604
Db |||||||
QY 545 AGACGGTGAGGACTGCAATTGCTCAATCTATCCCGGCGCAATAACGGGTACCGTATGG 604
Db |||||||
QY 605 CTTGGGATATGATGAAGTGGT 628
Db |||||||
QY 605 CTTGGGATATGATGAAGTGGT 628
Db |||||||
```

Search completed: December 20, 2003, 02:01:55
Job time : 2410.07 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 175.828 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-13

Perfect score: 636

Sequence: 1 ATGCTGGTGAAGCCATCGA.....TGATGAACGTGACTACTAATAG 636

Scoring table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	636	17	AAT12709
2	636	100.0	636	24	AAL48918
3	612.8	96.4	633	17	AAT12706
4	612.8	96.4	633	24	AAL48915
5	604	95.0	795	17	AAT12705
6	604	95.0	795	24	AAL48914
7	601.6	94.6	2082	24	AAL48939
8	601.6	94.6	2086	17	AAT12973

9	601.6	94.6	2433	17	AAT12974	HCV E1 construct H
10	590.6	92.9	2434	24	AAL48940	Hepatitis C virus
11	561.2	88.2	673	19	AAV42305	HCV core protein c
12	556.8	87.5	2187	19	ABA03491	Cuticle protein 1
13	556.8	87.5	2540	14	AAO43889	NANB hepatitis vir
14	556.8	87.5	2540	15	AAO63753	NANBHV genomic fra
15	556.8	87.5	9605	24	ABK91411	Hepatitis C virus
16	556.8	87.5	9605	24	ABK91424	Hepatitis C virus
17	556.8	87.5	9605	24	ABK91425	Hepatitis C virus
18	556.8	87.5	9605	24	ABK91426	Hepatitis C virus
19	556.8	87.5	9605	24	ABK91428	Hepatitis C virus
20	556.8	87.5	9605	24	ABK91429	Hepatitis C virus
21	556.8	87.5	9605	24	ABK91430	Hepatitis C virus
22	556.8	87.5	9605	24	ABK91431	Hepatitis C virus
23	556.8	87.5	9605	24	ABK91432	Hepatitis C virus
24	556.8	87.5	9605	24	ABK91433	Hepatitis C virus
25	556.8	87.5	9605	24	AAO25332	Hepatitis C virus
26	556.8	87.5	9608	24	ABK91427	Hepatitis C virus
27	556.8	87.5	11082	24	AAO25331	Hepatitis C virus
28	556.8	87.5	11076	21	AAA98965	Hepatitis C virus
29	555.2	87.3	1880	13	AAQ24467	NANB hepatitis vir
30	553.6	87.0	2540	13	AAQ29628	Hepatitis C virus
31	553.6	87.0	3461	15	AAO64068	Non-A, non-B hepat
32	553.6	87.0	3461	16	AAT30386	5'UTR/CORE/ENV/NSI
33	552	86.8	1251	13	AAQ26981	HCV gene 1. Hepat
34	552	86.8	1863	12	AAQ15363	Fragment of NANB h
35	552	86.8	3360	17	AAT03677	Hepatitis C genome
36	552	86.8	9413	16	AAT03960	Partial HCV non-st
37	552	86.8	9413	16	AAO81559	Hepatitis C virus
38	552	86.8	9413	16	AAO80498	DNA encoding HCV p
39	552	86.8	9413	24	AAO25517	Hepatitis C virus
40	552	86.8	9413	25	AAO49655	Hepatitis C virus
41	552	86.8	9413	25	AAO53723	Hepatitis C virus
42	551.4	86.7	1251	13	AAO25610	HCV in expression
43	551.4	86.7	1398	13	AAO22140	Hepatitis C virus
44	550.4	86.5	932	13	AAO20923	C10-E12 DNA fragme
45	550.4	86.5	1270	19	AAV60668	Fragment #1 isolat

ALIGNMENTS

RESULT 1
AAT12709
ID AAT12709 standard; DNA; 636 BP.

AC AAT12709;
XX
XX 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC117A.

KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.

OS Hepatitis C virus.

PN WO9604385-A2.

PD 15-FEB-1996.

PF 31-JUL-1995; 95WO-EP03031.

PR 29-JUL-1994; 94EP-0870132.

PR (INNO-) INNOGENETICS NV.

PI Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

XX Hepatitis C virus clone HCC11A E1 protein coding sequence.
DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX virucide; immunostimulant; vaccine; ds.
KW Hepatitis C virus.
XX WO200255548-A2.
XX 18-JUL-2002.
XX 11-JAN-2002; 2002WO-EP00219.
XX 11-JAN-2001; 2001US-260699P.
XX 30-AUG-2001; 2001US-315768P.
XX (INNO-) INNOGENETICS NV.
XX Maertens G, Bosman F, Buyse M;
XX WPI; 2002-599657/64.
XX P-PSDB; AAO18662.
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection -
XX
XX Example 2; Page 163-164; 243pp; English.
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX in vitro monitoring of HCV disease or prognosis. The response to
XX treatment of patients suffering from HCV infection. The present sequence
XX is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;
SQ
Query Match 96.4%; Score 612.8; DB 24; Length 633;
Best Local Similarity 98.1%; Pred. No. 1.3e-155;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATGCTGGGTAAAGCCATCGATACCTTACGTGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATGATACCTTACGTGCGGCTTCGCCGACCTCATGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGGCGCATGGGCTCCGGGTT 120
DB 61 CCGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGGCGCATGGGCTCCGGGTT 120
QY 121 CTGGAAGACGGCTGAATATGCAACAGGGAAATTTGCCGTGTTGCTCTTATCTTC 180
DB 121 CTGGAAGACGGCTGAATATGCAACAGGGAAATTTGCCGTGTTGCTCTTATCTTC 180
QY 181 CTCTTGCTTTACTGCTGCTGTCGACCAATTCAGCTTCAGGCTGCGCAACGCTG 240
DB 181 CTCTTGCTTTACTGCTGCTGTCGACCAATTCAGCTTCAGGCTGCGCAACGCTG 240
QY 241 TCCGGATGTACCATGTGACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCG 300
DB 241 TCCGGATGTACCATGTGACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTCCCTTGGTTCGGGAGAACACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGGTCCCTTGGTTCGGGAGAACACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTACCCCGACGCTCGGGCTAGGAAGCCAGCATCCCACTACAACA 420
DB 361 TGCTGGGTAGCGCTACCCCGACGCTCGGGCTAGGAAGCCAGCATCCCACTACAACA 420
QY 421 ATACGAGCGGCTCGATTGCTGCTGTTGGGGGGCTGCTTTCTGTTCCGCTATGACGTG 480
DB 421 ATACGAGCGGCTCGATTGCTGCTGTTGGGGGGCTGCTTTCTGTTCCGCTATGACGTG 480
QY 481 GGGGATCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTTCCATCTCGGCTCGCCG 540
DB 481 GGGGATCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTTCCATCTCGGCTCGCCG 540
QY 541 CATGAGACGGTGCAGGACTGCAATTCCTCAATCTATCCGGCCACATAACGGGTACCGT 600
DB 541 CATGAGACGGTGCAGGACTGCAATTCCTCAATCTATCCGGCCACATAACGGGTACCGT 600
QY 601 ATGGCTTGGGATATGATGATGAATGAACTGGTACTA 632
DB 601 ATGGCTTGGGATATGATGATGAATGAACTGGTAAATA 632
RESULT 5
AAT12705
ID AAT12705 standard; DNA; 795 BP.
XX AAT12705;
AC AAT12705;
XX 23-SEP-1996 (first entry)
DT
XX HCV E1 construct HCC110A.
DE
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
XX proteins - in presence of disulphide bond cleavage agent, to
XX produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or
XX the genotype of HCV. The new purification method preserves the


```
PT produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
PS Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;
SQ Query Match 94.6%; Score 601.6; DB 17; Length 2086;
Best Local Similarity 97.8%; Pred. No. 2e-152;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGTAAAGCCATCGATACCTTACGTGGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
DB |||||||
DB 5 TGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
QY 65 TCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTTCTGG 124
DB |||||||
DB 65 TCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTTCTGG 124
QY 125 AAGACGGCGTGAACACTATGCAACAGGGAATTTGCGTGGTGTCTTCTCTATCTTCTCT 184
DB |||||||
DB 125 AAGACGGCGTGAACACTATGCAACAGGGAATTTGCGGTTGCTTCTTCTATCTTCTCT 184
QY 185 TGGCTTTACTGCTGCTCTAAACATTCAGCTTTCGCTTACAGAGTGGCGAAGTGTCCG 244
DB |||||||
DB 185 TGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTTCGCTTATGAAGTGGCGAAGTGTCCG 244
QY 245 GGATGTACCATGTCACGACGACTGCTCAACTCAGCATTTGTATGAGGAGCGGACA 304
DB |||||||
DB 245 GGATGTACCATGTCACGACGACTGCTCAACTCAGCATTTGTATGAGGAGCGGACA 304
QY 305 TGATCATGCACACCCCCGGGTGCTGCTCCGCTTCGGGAGAACAACTCTTCCCGTGTCT 364
DB |||||||
DB 305 TGATCATGCACACCCCCGGGTGCTGCTCCGCTTCGGGAGAACAACTCTTCCCGTGTCT 364
QY 365 GGGTAGCGTCAACCCCAAGCTCGCGGTAGGAACGCCAGCATCCCACTPACAAACATAC 424
DB |||||||
DB 365 GGGTAGCGTCAACCCCAAGCTCGCGGTAGGAACGCCAGCATCCCACTPACAAACATAC 424
QY 425 GACGCCACGTCGATTTGCTGTTGGGGCGGCTCTTCTGTTCCGCTATGTAGTGGGG 484
DB |||||||
DB 425 GACGCCACGTCGATTTGCTGTTGGGGCGGCTCTTCTGTTCCGCTATGTAGTGGGG 484
QY 485 ATCTTCGCGGATCTGCTTCTCTGCTCTCCAGCTGTTTCAACATCTCGCTCGCGGCGCATG 544
DB |||||||
DB 485 ACCTTCGCGGATCTGCTTCTCTGCTCTCCAGCTGTTTCAACATCTCGCTTCGCGGCGCATG 544
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGG 604
DB |||||||
DB 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGG 604
QY 605 CTTGGGATATGATGATGAATGTTGT 628
DB |||||||
DB 605 CTTGGGATATGATGATGAATGTTGT 628
```

RESULT 9

```
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX
AC AAT12974;
XX
DT 25-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC166.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-BP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
(PINO-) INNOGENETICS NV.
XX
Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV.
XX
Claim 23; Fig 21; 146pp; English.
XX
AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
SQ Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
Query Match 94.6%; Score 601.6; DB 17; Length 2433;
Best Local Similarity 97.8%; Pred. No. 2.1e-152;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGTAAAGCCATCGATACCTTACGTGGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
DB |||||||
DB 356 TGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 415
QY 65 TCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTTCTGG 124
DB |||||||
DB 416 TCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTTCTGG 475
QY 125 AAGACGGCGTGAACACTATGCAACAGGGAATTTGCGTGGTGTCTTCTCTATCTTCTCT 184
```

476 AGGACGGGTGAATATGCAACAGGAAATTTGCCCGGTGCTCTTTCTATCTCTCT 535
185 TGGCTTTACTGCTCTTAACCATTCACAGCTTCGGCTTACGAGTGCGCAACGTGTCG 244
536 TGGCTTTGCTGCTGCTGACCGTTCCAGCTTCGGCTTATGAAGTGCGCAACGTGTCG 595
245 GGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGGACA 304
596 GGATGTACCATGTACGAAACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGGACA 655
305 TGATCATGACACCCCGGGTGGTCCCTGCTGCTGCGGAGAACAACTTTCCCGCTGCT 364
656 TGATCATGACACCCCGGGTGGTCCCTGCTGCTGCGGAGAACAACTTTCCCGCTGCT 715
365 GGGTAGCGGTACACCCCGAGCTCGCGGTAGGAACGCCAGCATCCCACTACAACATAC 424
716 GGGTAGCGGTACACCCCGAGCTCGAGCTAGGAACGCCAGCGTCCCAACGACAATAC 775
425 GAGCCACGTGCAATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTATGAGTGGGG 484
776 GAGCCACGTGCAATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTATGAGTGGGG 835
485 ATCTCTCGGATGCTCTTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCGGCGATG 544
836 ACCTCTGGGATGCTCTTCTGCTCTCCAGCTGTTCAACATCTCGCTCGCGGCGATG 895
545 AGACGGTGCAGGATGCAATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTATGAGTGG 604
896 AGACGGTGCAGGATGCAATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTATGAGTGG 955
605 CTTGGGATATGATGAACTGGT 628
956 CTTGGGATATGATGAACTGGT 979

RESULT 10
AAL48940
ID AAL48940 standard; DNA; 2434 BP.
XX
AC AAL48940;
XX
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
XX
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
PN WO20025548-A2.
XX
PD 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
PR 30-AUG-2001; 2001US-315768P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buysse M;
XX
XX WPI; 2002-599657/64.
DR P-PSDB; AAO18679.
XX
XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection -
XX
XX Example 2; Page 212-215; 243pp; English.

XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.

XX Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;

Query Match 92.9%; Score 590.6; DB 24; Length 2434;
Best Local Similarity 97.6%; Pred. No. 2e-149;
Matches 610; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 5 TGGGTAGGCGCATGATACCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 64
Db 356 TGGGTAAAGGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGCG 415
QY 65 TCGTCGGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGTTCGG 124
Db 416 TCGTCGGCGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGTTCGG 475
QY 125 AAGACGGGTGAATATGATGCAACAGGGAATTTGCCCTGCTGCTCTTCTATCTCTCTCT 184
Db 476 AGGACGGCGTGAATATGCAACAGGGAATTTGCCCGGTGCTCTTCTATCTCTCTCTCT 535
QY 185 TGGCTTTACTGTCCTG-TCTAACCATTTCCAGCTTCGCTTACGAGGTGCGCAAGTGTCC 243
Db 536 TGGCTTTGCTGCTGCTTTCACCGTTCCAGCTTCGCTTATGAAGTGCAGCGGTGCC 595
QY 244 GGGATGTACCATGTCCAGACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGCGGAC 303
Db 596 GGGATGTACCATGTCCAGACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGGAC 655
QY 304 ATGATCATGCACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
Db 656 ATGATCATGCACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGCTGC 715
QY 364 TGGGTAGCGCTCACCCCGAGCTCGGGCTAGGAACGCGAGCATCCCACTACAACAATA 423
Db 716 TGGGTAGCGCTCACCCCGAGCTCGCGAGTAGGAACGCGAGCATCCCACTACAACAATA 775
QY 424 CGACGCGAGTCAATTTGCTTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
Db 776 CGACGCGAGTCAATTTGCTTCGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 835
QY 484 GATCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTCCACCATCTCGCTCGCGCGCAT 543
Db 836 GACCTCTGCGGATCTGCTTCTCTCTCCAGCTGTTCCACCATCTCGCTCGCGCGCAT 895
QY 544 GAGACGGTGCAGGACTGCAATTCCTCAATCTATCCCGGCGACATAACGGGTACCGGTATG 603
Db 896 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAACGGGTACCGGTATG 955
QY 604 GCTTGGGATATGATGAACTGGT 628
Db 956 GCTTGGGATATGATGAACTGGT 980

RESULT 11
AAV42305
ID AAV42305 standard; cDNA; 673 BP.
XX
AC AAV42305;
XX
DT 26-OCT-1998 (first entry)
XX

DE HCV core protein cDNA insert of plasmid pA10.
 KW HBC; core protein; hepatitis D virus; L-HDAG; virus-like particle;
 KW infection; hepatitis B virus surface antigen; HBsAg; immunogen;
 KW vaccine; ds.
 XX Hepatitis C virus.
 PN WO9828004-A1.
 XX 02-JUL-1998.
 XX 24-DEC-1997; 97WO-AU000884.
 XX 24-DEC-1996; 96AU-0004341.
 XX (QUEB-) QUEBENS LAND DEPT HEALTH SAKZEWSKI VIRUS.
 PA Gowans EJ, MacNaughton TB;
 PI WPI; 1998-377411/32.
 XX P-PSDB; AAW62825.
 DR Virus-like particle for, e.g. treating microbial infection -
 XX comprises polypeptide from microorganism and sequence from Hepatitis
 PT D virus large protein, partially enveloped by Hepatitis B surface
 PT antigen
 XX
 PS Example; Fig 4; 72pp; English.
 XX
 CC This nucleotide comprises the hepatitis C virus (HCV) cDNA insert
 CC in plasmid pA10. This cDNA was obtained by PCR amplification of
 CC HCV cDNA using primers (see AAV38849-50) designed to amplify the
 CC HCV core gene. The PCR product was cloned into pBluescript KS to
 CC create pA10. It encodes a 224-amino acid polypeptide (see AAW62825).
 CC The HCV core protein can be used as an immunogen in novel fusion
 CC proteins (see AAW62657-59) that comprise HCV core protein and at
 CC least 19 amino acids (see AAW62827) of the C-terminal sequence of the
 CC large protein from hepatitis B virus (L-HDAG). In novel virus-like
 CC particles of the invention, a fusion protein immunogen is at least
 CC partially enveloped by hepatitis B surface antigen. The virus-like
 CC particle is used to ameliorate or protect against infections caused
 CC by hepatitis B virus and/or another microorganism, especially HCV.
 XX
 SQ Sequence 673 BP; 115 A; 208 C; 188 G; 162 T; 0 other;
 Query Match 88.2%; Score 561.2; DB 19; Length 673;
 Best Local Similarity 93.9%; Pred. No. 1.1e-141;
 Matches 584; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 7 GGTAAGCGCCATCGATACCTTACGTGCGGCTTCGCGGACCTCGTGGGTACATTCCGCTC 66
 DB 1 GGTAAGTTCATCGATACCTTACGTGCGGCTTCGCGGACCTCATGGGTACATTCCGCTC 60
 QY 67 GTCGGCGCCCTTAGGGGCGCTGCGAGGCGCTCGCGCATGGCTCGGGTTCGGAA 126
 DB 61 GTCGGCGCCCTTAGGGGCGCGCGCGAGGCGCTCGCGCATGGCTCGGGTTCGGAG 120
 QY 127 GACGGCGTGAACATATGCAACAGGAAATTCCTGGTGTCTTCTATCTTCCTCTTG 186
 DB 121 GACGGCGTGAACATATGCAACAGGAAATTCGCCGGTGTCTTCTATCTTCCTCTTG 180
 QY 187 GCTTTACTGTCTGTCTAAACCTTCAGCTTCGGCTTACGAGTGCACAGCTGTCGGG 246
 DB 181 GGTTCGTCTGTCTTTGACCATCCAGCTTCGCTTATGAAGTGGCAACGHTGCCGG 240
 QY 247 ATGTACCATGTACGAACGACTCTCCAACTCAAGCATTTGTATGAGGCGAGCGACATG 306
 DB 241 GTGTACCATGTACGAACGACTCTCCAAAGCAAGCATTTGTATGAGGCGAGCGACATG 300
 QY 307 ATCATGCACACCCCGGGTGCCTGCTGGTTCGGGAGAACAACTCTTCCCGCTGCTGG 366
 DB 301 ATCATGCACGTCCCGGGTGCCTGCTGGTTCGGGAGAACAACTCTTCCCGCTGCTGG 360

367 GTAGCGCTCACCCCAACGCTCGCGGTAGGAACGCGCAGCATCCCACTACAACAATACGA 426
 DB 361 GTAGCGCTCACCCCAACGCTTCGGCCAGGAACGCTAGCGTCCCTACTACGGCAATACGA 420
 QY 427 CGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTTGGGGAT 486
 DB 421 CGCCACGTCGATTTGCTCGTTGGGGCGGCTACTTTCTGTTCGCTATGTACGTTGGGGAT 480
 QY 487 CTCTCGGATCTGTCTTCTCTCCAGCTGTTCACCATCTCTCGCTCGCCGGCATGAG 546
 DB 481 CTCTCGGATCTGTCTTCTCTCGCCAGCTGTTCACCATCTCTCGCCGGCATGAG 540
 QY 547 AGGTGCGAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCT 606
 DB 541 AGGTGCGAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGCATGGCT 600
 QY 607 TGGGATATGATGATGAACCTGGT 628
 DB 601 TGGGATATGATGATGAACCTGGT 622

RESULT 12
 ABA03491
 ID ABA03491 standard; DNA; 2187 BP.
 XX
 AC ABA03491;
 XX
 DT 15-MAR-2002 (first entry)
 XX
 DE Cuticle protein 1 and 2 secreting hepatitis C virus related DNA #1.
 XX
 KW Cuticle protein 1; cuticle protein 2; hepatitis C virus; ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2187
 FT /*tag= a
 FT /product= "AAW47264"
 FT /partial
 FT /note= "no stop codon"
 XX
 PN KR97065713-A.
 XX
 PD 13-OCT-1997.
 XX
 PF 19-MAR-1996; 96KR-0007404.
 XX
 PR 19-MAR-1996; 96KR-0007404.
 XX
 PA (GLDS) LG CHEM LTD.
 XX
 PI Choo SH, Lee IH, Ryoo WS;
 XX
 DR WPI; 1998-492654/42.
 DR P-PSDB; AAW47264.
 XX
 XX Cuticle protein 1 and 2 secreting hepatitis C virus (Japanese)
 PT NoAbstract -
 XX
 PS Disclosure; Page 2-4; 7pp; Korean.
 XX
 CC The present invention relates to cuticle protein 1 and 2 secreting
 CC hepatitis C virus. The present sequence is a coding sequence
 CC provided in the exemplification of the invention.
 XX
 SQ Sequence 2187 BP; 406 A; 669 C; 631 G; 481 T; 0 other;
 Query Match 87.5%; Score 556.8; DB 19; Length 2187;
 Best Local Similarity 93.3%; Pred. No. 2.5e-140;
 Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

12-JUN-1990; 90JP-0153401.
08-NOV-1990; 90JP-0304405.
(NAKA/) NAKAMURA T.
WPI; 1993-199637/25.
P-PSDB; AAR38279.
Antigen related to non-A and non-B hepatitis virus - comprises
non-translation region comprising 340 - 341 mols. of nucleotides,
PT non-translation region comprising 1885 - 2551 mols. of
PT nucleotides including region 1,149 and, etc.
XX Claim 3; Page 19-20; 73pp; Japanese.
PS The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC The polypeptide it encodes may be used in a system for detecting
CC NANB hepatitis. This method is highly specific and sensitive, and
CC can detect NANB hepatitis virus which could not be detected by
CC conventional methods.
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
SQ Query Match 87.5%; Score 556.8; DB 14; Length 2540;
Best Local Similarity 93.3%; Pred. No. 2.7e-140;
Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 5 TGGGTAAGCCATCGATACCCCTTACGTCGGGCTTCGCGACCTCGTGGGGTACATTCGCG 64
Db |||||
QY 356 TGGGTAAGGTCATCATATACCCCTTACATGCGGCTTCGCCGATCTCATGGGTATATTCGCG 415
Db |||||
QY 65 TCGTCGGCGCCCTAGGGGCGCTGCGCAGGCGCTGCGCATGCGCTCGGGTTCGG 124
Db |||||
QY 416 TCGTCGGCGCCCTAGGGGCGCTGCGCAGGCGCTTGGCACACGCTGCGGGTTCGG 475
Db |||||
QY 125 AAGACGGGTGAATGCAAGGAAATTCGCTGGTTCCTCTCTCTCTCTCTCTCTCTCT 184
Db |||||
QY 476 AAGACGGGTGAATGCAAGGAAATTCGCTGGTTCCTCTCTCTCTCTCTCTCTCTCT 535
Db |||||
QY 185 TGGCTTTACTGCTCTGTAACCATTCAGCTTCGCTTACGAGGTGCGCAACGCTGCG 244
Db |||||
QY 536 TGGCTTTGCTGCTGCTGTTGACCATTCAGCTTCGCTTATGAGGTGCGCAACGCTGCG 595
Db |||||
QY 245 GGATGTACCATGTACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGGACA 304
Db |||||
QY 596 GGATATACCATGTACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGGACA 555
Db |||||
QY 305 TGATCATGCACACCCCGGCTGCGCTTCGCTTCGGGAGAACACTCTTCCCGTGTCT 364
Db |||||
QY 656 TGATCATGCATATCCCGGCTGCGCTTCGCTTCGGGAGAACACTCTTCCCGTGTCT 715
Db |||||
QY 365 GGGTAGCGCTCACCCCGCAGCTCGCGCTAGGAACGCGCAGCATCCCGCATCAACATAC 424
Db |||||
QY 716 GGGTAGCGCTCACCCCGCAGCTCGCGCTAGGAACGCGCTCCCGCATCAACATAC 775
Db |||||
QY 425 GACGCCAGCTGATTCGCTGGGGGCGCTGCTTCTGCTTCGCTATGTACGTGGGG 484
Db |||||
QY 776 GACGCCAGCTGATTCGCTGGGGGCGCTGCTTCTGCTTCGCTATGTACGTGGGG 835
Db |||||
QY 485 ATCTCTGGGATCTGCTTCCTGCTCCAGCTGTTTCCATCTCGCTTCGCTTCGCTTCG 544
Db |||||
QY 836 ATCTCTGGGATCTGCTTCCTGCTCCAGCTGTTTCCATCTCGCTTCGCTTCGCTTCG 895
Db |||||
QY 545 AGACGGTGCAGGACTGCAATTCATCTATCCCGGCGCATTAACGGGTACCGGTATGG 604
Db |||||
QY 896 AGACAGTGCAGGACTGCAATTCATCTATCCCGGCGCATTTATATAGGTCACCGATGG 955
Db |||||
QY 605 CTTGGGATATGATGAACCTGGT 628
Db |||||
QY 956 CTTGGGATATGATGAACCTGGT 979
Db |||||

RESULT 13
AAQ43889 standard; cDNA to mRNA; 2540 BP.
AC AAQ43889;
XX

21-OCT-1993 (first entry)

NANB hepatitis virus polynucleotide N-2540-2.

XX Non-A, non-B; virus; polymerase chain reaction; detection;
KW sensitive; specific; HCV; NANBH; ss.
XX

Non-A, non-B hepatitis virus.

Key Location/Qualifiers

FT CDS 342..2540

FT /tag= a

FT 1..341

FT /tag= b

FT /note= "from 5' terminal of NANBH virus RNA"

XX JF05091884-A.

XX 16-APR-1993.

XX 10-APR-1991; 91JP-0196175.

XX

PR 12-JUN-1990; 90JP-0153401.
PR 08-NOV-1990; 90JP-0304405.
XX (NAKA/) NAKAMURA T.
XX WPI; 1993-199637/25.
XX P-PSDB; AAR38279.
XX Antigen related to non-A and non-B hepatitis virus - comprises
XX non-translation region comprising 340 - 341 mols. of nucleotides,
XX PT non-translation region comprising 1885 - 2551 mols. of
XX PT nucleotides including region 1,149 and, etc.
XX Claim 3; Page 19-20; 73pp; Japanese.
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
XX which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
XX The polypeptide it encodes may be used in a system for detecting
XX NANB hepatitis. This method is highly specific and sensitive, and
XX can detect NANB hepatitis virus which could not be detected by
XX conventional methods.
XX Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;
SQ Query Match 87.5%; Score 556.8; DB 14; Length 2540;
Best Local Similarity 93.3%; Pred. No. 2.7e-140;
Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 5 TGGGTAAGCCATCGATACCCCTTACGTCGGGCTTCGCGACCTCGTGGGGTACATTCGCG 64
Db |||||
QY 65 TCGTCGGCGCCCTAGGGGCGCTGCGCAGGCGCTGCGCATGCGCTCGGGTTCGG 124
Db |||||
QY 757 TCGTCGGCGCCCTAGGGGCGCTGCGCAGGCGCTTGGCACACGCTGCTCGGGTTCGG 816
Db |||||
QY 125 AAGACGGGTGAATGCAAGGAAATTCGCTGGTTCCTCTCTCTCTCTCTCTCTCTCT 184
Db |||||
QY 817 AAGACGGGTGAATGCAAGGAAATTCGCTGGTTCCTCTCTCTCTCTCTCTCTCTCT 876
Db |||||
QY 185 TGGCTTTACTGCTCTGTAACCATTCAGCTTCGCTTACGAGGTGCGCAACGCTGCG 244
Db |||||
QY 877 TGGCTTTGCTGCTGCTGTTGACCATTCGCTTATGAGGTGCGCAACGCTGCG 936
Db |||||
QY 245 GGATGTACCATGTACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGGACA 304
Db |||||
QY 937 GGATATACCATGTACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGGACA 996
Db |||||
QY 305 TGATCATGCACACCCCGGCTGCGCTTCGCTTCGGGAGAACACTCTTCCCGTGTCT 364
Db |||||
QY 997 TGATCATGCATATCCCGGCTGCGCTTCGCTTCGGGAGAACACTCTTCCCGTGTCT 1056
Db |||||
QY 365 GGGTAGCGCTCACCCCGCAGCTCGCGCTAGGAACGCGCAGCATCCCGCATCAACATAC 424
Db |||||
QY 1057 GGGTAGCGCTCACCCCGCAGCTCGCGCTAGGAACGCGCAGCATCCCGCATCAACATAC 1116
Db |||||
QY 425 GACGCCAGCTGATTCGCTGGGGGCGCTGCTTCTGCTTCGCTATGTACGTGGGG 484
Db |||||
QY 1117 GACGCCAGCTGATTCGCTGGGGGCGCTGCTTCTGCTTCGCTATGTACGTGGGG 1176
Db |||||
QY 485 ATCTCTGGGATCTGCTTCCTGCTTCGAGCTGTTTCCATCTCGCTTCGCTTCGCTTCG 544
Db |||||
QY 1177 ATCTCTGGGATCTGCTTCCTGCTTCGAGCTGTTTCCATCTCGCTTCGCTTCGCGCATG 1236
Db |||||
QY 545 AAGCGGTGCAGGACTGCAATTCATCTATCCCGGCGCATTAACGGGTACCGGTATGG 604
Db |||||
QY 1237 AAGCAGTGCAGGACTGCAATTCATCTATCCCGGCGCATTTATATAGGTCACCGATGG 1296
Db |||||
QY 605 CTTGGGATATGATGAACCTGGT 628
Db |||||
QY 1297 CTTGGGATATGATGAACCTGGT 1320
Db |||||

```
RESULT 14
AAQ63753
ID AAQ63753 standard; cDNA to mRNA; 2540 BP.
XX
AC AAQ63753;
XX
DT 30-JAN-1995 (first entry)
XX
DE NANBHV genomic fragment #2.
XX
XX Polymerase chain reaction; PCR; primer; amplify; detection; NANBHV;
KW non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
XX
OS Synthetic.
XX
PN JP06125777-A.
XX
PD 10-MAY-1994.
XX
PF 20-JUN-1991; 91JP-0247120.
XX
PR 20-JUN-1991; 91JP-0247120.
XX
PA (NAKA/) NAKAMURA T.
XX
DR WPI; 1994-187937/23.
XX
XX Oligonucleotide primer pairs specific for non-A, non-B hepatitis
PT virus - used for high sensitivity detection of non-A non-B (NANB)
PT hepatitis virus
XX
PS Disclosure; Page 24-25; 25pp; Japanese.
XX
CC The sequences given in AAQ63752-53 represent fragments of the non-A,
CC non-B hepatitis virus (NANBH) genome. These fragments were amplified
CC using the primers given in AAQ63732-51. These primers were used in the
CC detection of NANBH. The primers are based on the 5'-terminal region and
CC the core protein coding region. The method allows highly sensitive
CC detection of NANBH.
XX
SQ Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

Query Match 87.5%; Score 556.8; DB 15; Length 2540;
Best Local Similarity 93.3%; Pred. No. 2.7e-140;
Matches 582; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 5 TGGGTAAGGCCATCGATACCTTACCTGGGGCTTCGCCGACCTCGGGGTACATTCGCG 64
DB 697 TGGGTAAGGTCATCGATACCTTACATGCGGCTTCGCCGATCTCATGGGGTATATTCGCG 756
QY 65 TCGTGGGCGCCCTTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 124
DB 757 TCGTGGGCGCCCTTAGGGGGCGCTGCAGGGCCCTGGCACACGGGTGCGGGTTCTGG 816
QY 125 AAGACGGCGTGAATGCAACAGGGAATTTGCTGGTGTCTTCTCTATCTCTCTCT 184
DB 817 AGACGGCGTGAATGCAACAGGGAATTTGCTGGTGTCTTCTCTATCTCTCTCTCT 876
QY 185 TGGCTTTACTGCTCTGTAACCATTCAGCTTCCGCTTACGAGGTGCGCAAGCTGTCGG 244
DB 877 TGGCTTTGCTGCTGTTTACCATCCGAGCTTCCGCTTATGAAGTGGCAAGCTGTCGG 936
QY 245 GGATGTACCATGTACGAACAGCACTGTCCAACTCAAGCATTTGTGTATGAGGCGCGGACA 304
DB 937 GGATATACCATGTACGAACAGCACTGTCCAACTCAAGCATTTGTGTATGAGGCGCGGACA 996
QY 305 TGATCATGACACCCCGGGTGGTCCCTCGCTTCCGGAGAACAACTCTTCCCGTGTCT 364
DB 997 TGATCATGATACCTCCCGGGTGGTCCCTCGCTTCCGGAGAACAACTCTTCCCGTGTCT 1056
QY 365 GGGTAGCGCTCACCCCGCGCTCGGCTAGGACGCGAGCATCCCGCACTACAAATAC 424
DB 1057 GGGTAGCGCTCACTCCCAAGCTCGCGGCGAGGAATCCAGCGTCCCGCACTACAAATAC 1116
```

```
QY 425 GACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCGGGG 484
DB 1117 GACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGATGCGGGG 1176
QY 485 ATCTCTGCGGATCTGTTCTTCTGCTCTCCAGCTGTTCCACATCTCGCCTCGCGGCATG 544
DB 1177 ATCTCTGCGGATCTGTTCTTCTGCTCTCCAGCTGTTCCACATCTCGCCTCGCGGCATG 1236
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGG 604
DB 1237 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCATTTATCAGGTACCGCATGG 1296
QY 605 CTTGGGATATGATGATGAAGTGT 628
DB 1297 CTTGGGATATGATGATGAAGTGT 1320

RESULT 15
ABK91411
ID ABK91411 standard; DNA; 9605 BP.
XX
AC ABK91411;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
XX
PN WO200259321-A2.
XX
XX 01-AUG-2002.
XX
PF 16-JAN-2002; 2002WO-BP00526.
XX
PR 23-JAN-2001; 2001US-263479P.
XX
PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI De Francesco R, Migliaccio G, Paonessa G;
XX
DR WPI; 2002-599793/64.
XX
DR P-PSDB; ABG32451.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression
XX
PS Claim 9; Page 36-39; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1636.15 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-13
Perfect score: 636
Sequence: 1 ATGCTGGTAAAGCCATCGA.....TGATGAACGTGTAATAATAG 636

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pin.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_pbg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	75.4	11.9	488	9	AV755731
C 2	61.	9.6	492	9	AV758366
C 3	42	6.6	1201	13	BX356664
C 4	41.6	6.5	534	14	CD040840

CD040840 psHB036XB

C 5	41.4	6.5	1270	12	BG968359
C 6	41	6.4	1201	13	BX381961
C 7	40.4	6.4	1097	9	AL580130
C 8	40.2	6.3	1201	9	AL513886
C 9	39.6	6.2	925	29	CNS0091P
C 10	39	6.1	359	12	BJ252669
C 11	39	6.1	375	12	BJ246716
C 12	39	6.1	621	14	CAB16001
C 13	39	6.1	856	29	BZ578381
C 14	39	6.1	872	29	BZ555011
C 15	39	6.1	1201	13	BX460099
C 16	38.8	6.1	533	29	CC010084
C 17	38.8	6.1	639	29	CC405164
C 18	38.8	6.1	826	29	BZ736582
C 19	38.8	6.1	895	29	CC359028
C 20	38.8	6.1	925	29	CC359026
C 21	38.8	6.1	940	29	CC010085
C 22	38.8	6.1	951	29	CC405167
C 23	38.6	6.1	399	9	AV638521
C 24	38.6	6.1	434	9	AV637507
C 25	38.6	6.1	440	9	AV637983
C 26	38.6	6.1	450	9	AV637259
C 27	38.6	6.1	451	9	AV637328
C 28	38.6	6.1	451	9	AV637643
C 29	38.6	6.1	453	9	AV634724
C 30	38.6	6.1	454	9	AV637050
C 31	38.6	6.1	456	9	AV635382
C 32	38.6	6.1	473	9	AV632765
C 33	38.6	6.1	481	9	AV635503
C 34	38.6	6.1	485	9	AV632811
C 35	38.6	6.1	506	9	AV392445
C 36	38.6	6.1	508	9	AV634095
C 37	38.6	6.1	526	9	AV641895
C 38	38.6	6.1	533	9	AV638125
C 39	38.6	6.1	537	9	AV632335
C 40	38.6	6.1	588	9	AV387329
C 41	38.6	6.1	945	29	CNS050MQ
C 42	38.4	6.0	525	10	BE337089
C 43	38.4	6.0	671	12	BI723733
C 44	38.2	6.0	431	9	AV639153
C 45	38.2	6.0	501	9	AV638474

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

Location/Qualifiers

```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 116 a 134 c 137 g 97 t
ORIGIN
Query Match 11.9%; Score 75.4; DB 9; Length 488;
Best Local Similarity 70.3%; Pred. No. 3.6e-09; Indels 4; Gaps 2;
Matches 130; Conservative 0; Mismatches 51;

QY 445 GTTGGGGCGCTCTTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTCTCTTC 504
D 472 GTGGTGTACACCTCGCTCTGCTCAGCTCTCTAGTGTGGGACCTCTCGACGGAGTGATG 413
QY 505 CTGCTCTCCAGCTGTTCACCTCTCGCTCGCGGCATGAGACGTCGAGCACTGCAAT 564
D 412 CTTGCACTCAGCTG---ATCATCTGCGCTCAGCACCATGAGTTGTGATGAATGCAAC 356
QY 565 TGCTCATCTATCCCGCCACATAACGGGTCACCGTATG-GCTTGGGATATGATGAA 623
D 355 TGCTCCATCTATCTCGCGGCATCATCTGGACACCGTATGACATGGACATGATGAA 296
QY 624 CTGGT 628
D 295 CTGGT 291

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 124 a 128 c 125 g 112 t
ORIGIN
source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 116 a 134 c 137 g 97 t
ORIGIN
Query Match 11.9%; Score 75.4; DB 9; Length 488;
Best Local Similarity 70.3%; Pred. No. 3.6e-09; Indels 4; Gaps 2;
Matches 130; Conservative 0; Mismatches 51;

QY 445 GTTGGGGCGCTCTTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTCTCTTC 504
D 472 GTGGTGTACACCTCGCTCTGCTCAGCTCTCTAGTGTGGGACCTCTCGACGGAGTGATG 413
QY 505 CTGCTCTCCAGCTGTTCACCTCTCGCTCGCGGCATGAGACGTCGAGCACTGCAAT 564
D 412 CTTGCACTCAGCTG---ATCATCTGCGCTCAGCACCATGAGTTGTGATGAATGCAAC 356
QY 565 TGCTCATCTATCCCGCCACATAACGGGTCACCGTATG-GCTTGGGATATGATGAA 623
D 355 TGCTCCATCTATCTCGCGGCATCATCTGGACACCGTATGACATGGACATGATGAA 296
QY 624 CTGGT 628
D 295 CTGGT 291

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 124 a 128 c 125 g 112 t
ORIGIN

```

```

Query Match 9.6%; Score 61; DB 9; Length 492;
Best Local Similarity 68.0%; Pred. No. 2.1e-05; Indels 4; Gaps 2;
Matches 115; Conservative 0; Mismatches 50;

QY 461 TCTGTTCCGCTATGTACGTGGGGGATCTCTGGGGATCTCTCTCTCTCTCTCTCTCT 520
D 457 TGTGATCAGCTCTACTACGTGTGGACCTCTGCTTGGGGTATCGCTTGCAGCCCACTG- 399
QY 521 TCACCATCTCTGCTCCCGGCATGAGACGTCGAGACGTCGCAATTCCTCAATCTATCCG 580
D 398 --ATTATCTCTCAGCAGCAACATTGCTTTGTGCAAGAATGCAACTGCTCTATTATCTG 341
QY 581 GCACATACGGGTC-ACCGTATGGCTTGGGATATGATGATGAACCTGGT 628
D 340 GCTGCATCACTGGACATACAGTATGGCATAGGCTATGATGATGAACCTGGT 292

RESULT 3
BX356664 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSODI015YB03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1 GI:30378083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI015CA02NP1.
FEATURES
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI015YB03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 6.6%; Score 42; DB 13; Length 1201;
Best Local Similarity 10.7%; Pred. No. 2.6; Indels 0; Gaps 0;
Matches 57; Conservative 245; Mismatches 232;

QY 23 CCTTACGTGCGCTTTCGCGACCTCGTGGGGTACATCCGCTCGTCGGCGCCCTAG 82
D 665 SSBTBSYTBSSSTSSSSSBTTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 724
QY 83 GGGCGCTGCCAGGCGCCCTGGCCATGCGGCTTCGGAGACGGCGTGAACATG 142
D 725 SSTBSBTTCTTTKSSBSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 784
QY 143 CAACAGGGAATTGCGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 202
D 785 SCSTSSBSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 844

```


us-09-899-303a-13.rst

Mon Dec 22 13:28:36 2003

Db 1125 AGTGGCCCTTGGTGGTCTCTCTCTGGCGGTGACATCTCTGGCGGTCTGGCCCTC 1066

QY 141 TGAACAGAGGAATTTGGCTGGTCTCTTTCTCTATCTCTCTTGGCTTTACTGTCTG 200

Db 1065 TGCCTGACATCTGTCTCTGGTGGCATACTCTCCGCTGTCCCGGAGTGGTCTTGTCCCG 1006

QY 201 TCTAACCACTTCAGCTTCGGCTTACAGGTGGCAACGCTGCGGGATGTACCATGTCAAC 260

Db 1005 TCCATCTCTCCGCGCATCTCTGGTGGGTCCGATCTCCGCTGTCCCGGAGTGGTCTTGTCCCG 946

QY 261 GAACGACTCTCCAACTCAAGCATTTGTGATGAGGAGGAGCATGATCATGCACACCC 320

Db 945 CATCCCTCTGTGATCTTGTATGGGCTCTCTCCGCTGTCCCGGAGTGGTCTTGTCCCG 886

QY 321 CGGGTGGTGGTCTCTGGTGGGAGCAACTCTCTCCGCTGTCTGGTGGGCTCACCCC 380

Db 885 CGGGTGGTGGTCTCTGGTGGGAGCAACTCTCTCCGCTGTCTGGTGGGCTCACCCC 826

QY 381 CACGCTCGGGCTAGGAAGCCAGCATCCCC 411

Db 825 TTGCTTGGCTCTTGGGTACCTCTTCCGC 795

RESULT 6

BX381961/c 1201 bp mRNA linear EST 08-MAY-2003

LOCUS

DEFINITION BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI072YF05 3-PRIME, mRNA sequence.

ACCESSION BX381961

VERSION BX381961.1 GI:30453007

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1201)

TITLE Li, W.B., Gruber, C., Jesses, J. and Polayes, D.

JOURNAL Full-length cDNA libraries and normalization

COMMENT Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI072CC03NP1.

FEATURES

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI072YF05"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 95 a 191 c 115 g 55 t 745 others

ORIGIN

Query Match 6.4%; Score 41; DB 13; Length 1201;

Best Local Similarity 3.3%; Pred.No. 4.8;

Matches 19; Conservative 189; Mismatches 369; Indels 0; Gaps 0;

QY 30 GTGGCGCTTCGGCACTCGTGGGGTACATTCGGCTCGTGGCGCCCGCTAGGGGGCGC 89

Db 970 GRWVKBMGKMKMKKBMKBNKKTMMWMBKMKMKMKMKMKMKMKMKMKMKMKMKMKMK 911

QY 90 TGGCAGGGCCCTGGCGCATGGCTCGGGTCTCGGAGACGGCGTGAACATGCAACAGG 149

Db 910 MMKMGKMK 851

QY 150 GAATTTGGCTGGTGGTCTTTCTCTATCTCTCTTGGCTTTACTGTCTCTGTCTAACCAT 209

Db 850 KMMWVKRMKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMKKMK 791

QY 210 TCCAGCTTCGGCTTACGAGTGGCAACGCTGTCGGGATGTACCATGTTCAGAAACGATG 269

Db 790 MSKKKNN 731

QY 270 CTCCAACTCAAGCAATTTGTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGCT 329

Db 730 KMMWNN 671

QY 330 GCCCTGGTTCGGGAGCAACACTCTTCCCGCTGCTGGGTAGCGCTCACCCCGCCTCGC 389

Db 670 NNN 611

QY 390 GGCTAGGAAAGCCAGCATCCCACTACAACTACGACGCGCATGCTGCTTGTCTCTGGT 449

Db 610 NNN 551

QY 450 GCGCGCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTCGGATCTGCTCTCTCTCGT 509

Db 550 MNKKKNN 491

QY 510 CTCCAGCTGTTTCAACCATCTCGCTCGCGGCGCATGACGCGTGCAGGACTGCAATGCTC 569

Db 490 NNN 431

QY 570 AATCTATCCCGGCGCACATAACGGGTACCGGTATGGCT 606

Db 430 MNKKKNN 394

RESULT 7

AL580130/c 1097 bp mRNA linear EST 01-JUN-2003

LOCUS

DEFINITION AL580130 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED Homo sapiens cDNA clone CSODJ001YK17 3-PRIME, mRNA sequence.

ACCESSION AL580130

VERSION AL580130.2 GI:31318409

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1097)

TITLE Li, W.B., Gruber, C., Jesses, J. and Polayes, D.

JOURNAL Full-length cDNA libraries and normalization

COMMENT Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODJ001AF09NP1.

FEATURES

Location/Qualifiers

1..1097

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODJ001YK17"

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/cell_line="JURKAT"

/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"


```
source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 6.2%; Score 39.6; DB 29; Length 925;
Best Local Similarity 12.1%; Pred. No. 10;
Matches 45; Conservative 164; Mismatches 163; Indels 0; Gaps 0;

Qy 2 TGTGGTAAAGCCATGATACCTTACGTGGGCTTCGCCACCTCGTGGGGTACATTC 61
Db 553 TTSSGGYGGSGSGBSCSCSSCSGSSSCBCCBCCSCSSYCCSSBSKSSKSS 612

Qy 62 CGCTCGTCGGCGCCCTAGGGCGCTGCAGAGGCCCTGGCGCATGCGTCCGGTTC 121
Db 613 TSBSGCCSSKSVCGTSCSSSSSSSSSTSSSTSSSTSSSTSSSSSSSSSYTTS 672

Qy 122 TGAAGAGCGCGTAACTATGCAACAGGGAATTGCTGCTCTTTCTCTATCTTCC 181
Db 673 KSTASGSGWSAGSGSGSTGTSSTSSSSSTSSSSVSSGSKSBTSBSGSSBSGSS 732

Qy 182 TCTGGCTTTACTGCTCTGCTTAACATTCAGCTTCGCTACGAGGTGCGCAACGTG 241
Db 733 SSSTSSBSCTSSSSSSSSSYSSSTSCCTCCCSYSSSTSSSTSSSTSSSTSSSV 792

Qy 242 CCGGATGTACCATGTCCAGAACAGCTGCTCAACTCAAGCATGTTGTATGAGCGCG 301
Db 793 GTSSSSDSTSTCCSCCCYMCCTCCSTYBMCYTSTSCGSSSSSGKGGVTKCGCGGSS 852

Qy 302 ACATGATCATGCACACCCCGGTGCTGCGCTCGGAGAACAACTCTTCCCGCT 361
Db 853 TNGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSV 912

Qy 362 GCTGGTAGCGC 373
Db 913 SSGSGSGSGSVS 924

RESULT 10
BU252669/c 359 bp mRNA linear EST 05-APR-2002
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
BU252669.1 GI:20061830
VERSION
EST.
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
/clone 107 c 110 g 77 t

BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN

Query Match 6.1%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 11;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 24 CCTTACGTGGGCTTCGCGACCTCGTGGGGTACATTCGCTCGTGGCGCCCTTAGG 83
Db 36 CTTCAAGTGCAACAGCGCGCTCTGGAAGGCGCTCAGGGGGTTCAGCGCGTCCGTCGG 95

Qy 84 GGGCGCTGCCAGGCGCTGGCGCATGGCTCCGGTTCGGAAGACGGGTGAACATG 142
Db 96 GGACGCGCGAGCGCCCTGGGCGGAGGACGTGCGAGTCTGCGGTGACGTGCCCAAG 154

RESULT 11
BU246716 375 bp mRNA linear EST 05-APR-2002
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
BU246716
VERSION
EST.
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
/clone 107 c 110 g 77 t

BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN

Query Match 6.1%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 11;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 24 CCTTACGTGGGCTTCGCGACCTCGTGGGGTACATTCGCTCGTGGCGCCCTTAGG 83
Db 297 CTTCAAGTGCAACAGCGCGCTCTGGAAGGCGCTCAGGGGGTTCAGCGCGTCCGTCGG 238

Qy 84 GGGCGCTGCCAGGCGCTGGCGCATGGCTCCGGTTCGGAAGACGGGTGAACATG 142
Db 237 GGACGCGCGAGCGCCCTGGGCGGAGGACGTGCGAGTCTGCGGTGACGTGCCCAAG 179

RESULT 12
CA816001 621 bp mRNA linear EST 09-DEC-2002
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
CA816001
VERSION
EST.
KEYWORDS
Triticum aestivum (bread wheat)
SOURCE
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
/clone 107 c 110 g 77 t

BASE COUNT 70 a 115 c 107 g 67 t
ORIGIN
```

DEFINITION CAL2EI1011VF_E04 Cabernet Sauvignon Leaf - CAL2EI Vitis vinifera
 CDNA clone CAL2EI1011VF_E04 5', mRNA sequence.
 ACCESSION CAB16001
 VERSION CAB16001.1 GI:26264938
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; Vitaceae; Vitis.
 REFERENCE 1 (bases 1 to 621)
 AUTHORS Goes da Silva,F., Lim,H., Iandolino,A., Baek,J., Jones,K., Walker
 ,M.A. and Cook,D.R.
 TITLE Transcriptional responses of Vitis vinifera to infection by the
 bacterial pathogen Xylella fastidiosa
 JOURNAL Unpublished
 COMMENT Contact: Doug Cook
 CAES Genome Facility
 UC Davis Department of Plant Pathology
 1 Shields Ave., Davis, CA 95616, USA
 Tel: 530 754 8561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seq primer: GTTATCAGTCGACGGTACC.
 Location/Qualifiers

FEATURES
 source

1..621
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAL2EI1011VF_E04"
 /sex="hermaphrodite"
 /dev_stage="Mid-season leaf material"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Leaf - CAL2EI"
 /note="Organ: Leaf; Vector: pDNR; Site 1: Sfil; Site 2:
 Sfil; CAL2EI is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on July 25, 2001, in
 Napa Valley, California, and represent leaves in
 mid-season development. These leaves were verified to be
 infected with the bacterial pathogen, Xylella fastidiosa,
 based on a diagnostic assay using PCR and Xylella-specific
 primer pairs. The plants were asymptomatic at the time of
 collection, but later developed symptoms. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACAGTGGCATTACGCCGGG-3' and
 5'-ATTCTAGCGCGAGCGGCCACATG-dT(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."
 BASE COUNT 146 a 152 c 143 g 180 t

Query Match 6.1%; Score 39; DB 14; Length 621;
 Best Local Similarity 50.3%; Pred. No. 13;
 Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 264 CGACTGCTCCAACTCAAGCATTTGTATGAGGAGGAGGATCATGCACACCCCGG 323
 DB 397 CGACTTGTCTTCTATGACGGGTAGCGGAGGAGGACTCCGAGTCTCCGAG 456
 QY 324 GTGCGTGCCCTCGTTTCGGGAGAACATCTTTCCCGCTGCTGGTAGCGCTCACCCCCAC 383
 DB 457 GTGCTCGAGATGCTTGTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516
 QY 384 GTCGCGGTAGGAGCCAGCATTCCTCCACTACAAATACGACGCGCATGCTGCT 443
 DB 517 ACACCTTCGGTGTTCGGCAGCGTAGCCGCTGACACCATCGAAGCATAGCTGGTTCGG 576
 QY 444 CGTTGGGCGG 454
 DB 577 AGTAGTGGGG 587

RESULT 13
 BZ578381
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BZ578381 856 bp DNA linear GSS 17-DEC-2002
 msh2 5817.x1 msh Pseudomonas aeruginosa genomic clone msh2_5817,
 genomic survey sequence.
 BZ578381 GI:27213442
 GSS.
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 856)
 Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol., (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers

FEATURES
 source

1..856
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="MSH"
 /db_xref="taxon:287"
 /clone="msh2 5817"
 /clone_lib="msh"
 /note="Environmental isolate. Whole genomic shotgun
 library."
 BASE COUNT 147 a 274 c 279 g 154 t 2 others

Query Match 6.1%; Score 39; DB 29; Length 856;
 Best Local Similarity 58.0%; Pred. No. 14;
 Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 287 TGTATGAGCAGCGGACATGATGCACACCCCGGGTGGCGCTCGCTTCGGGAGA 346
 DB 253 TGTGGCGCGCGGACAGCCATCGCAGCGCGTGGCGGGTTCTGCTCGACGAGC 312
 QY 347 ACACTCTTCCCGTGTGGGTAGCGCTACCCCGCGCTCGCGGTAGGACGCCAGC 405
 DB 313 CCCACTCGGTGGCAGCAGGGTGGCGCTCGCTCGCTCGGTAGAGGTCCAGC 371

RESULT 14
 BZ555011
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BZ555011 872 bp DNA linear GSS 17-DEC-2002
 pacs1-60_4970.x1 pacs1-60 Pseudomonas aeruginosa genomic clone
 pacs1-60_4970, genomic survey sequence.
 BZ555011 GI:27163410
 GSS.
 Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 872)
 Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol., (2002) In press
 Contact: Chris K. Raymond
 Genome Center

Mon Dec 22 13:28:36 2003

University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .872
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone_lib="pac1-60 4970"
/notes="clinical isolate 1-60 whole genomic shotgun library."
152 a 262 c 284 g 173 t 1 others

BASE COUNT
ORIGIN

Query Match 6.1%; Score 39; DB 29; Length 872;
Best Local Similarity 58.0%; Pred. No. 15; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
Qy 287 TGTATGAGCGGACATGATGACACACCCCGGGTGGTGGCTTGGGAGA 346
Db 212 TGTGGCGCGGGACACGCCCATGCGAGCGGTGGCGGGTCTGTCGACGAGC 271
Qy 347 ACAACTTCCCGTGTGGTAGCGTACCCCGAGCTCGCGGTAGGAGCGCCAGC 405
Db 272 CCACACTCGTGGCAGCGGTGGCGCTTGGCTTGGCGGCGGTAGAGGTCCAGC 330

RESULT 15

EX460099/c
LOCUS
DEFINITION BX460099 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF010YP12 5-PRIME, mRNA sequence.

ACCESSION BX460099.1 GI:31029174

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li W.B., Gruber C., Jessee J. and Polayes D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DF010DH06QP1.

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DF010YP12"

/tissue_type="FETAL BRAIN"

/dev_stage="fetal"

/clone_lib="Homo sapiens FETAL BRAIN"

/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

188 a 292 c 194 g 110 t 417 others

BASE COUNT
ORIGIN

Query Match 6.1%; Score 39; DB 13; Length 1201;
Best Local Similarity 14.1%; Pred. No. 16; Indels 5; Gaps 2;
Matches 56; Conservative 174; Mismatches 162; Indels 5; Gaps 2;
Qy 67 GTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGGCGCATGGCGTTCGGGTTCTGGAA 126
Db 1185 KTSMGSGNGMMVMVMGKGKGMGSMGSMGGGKKACMMVKKMGKKMDGKMMKKKM 1126
Qy 127 GACGGCGTGAACATATGCAACAGGGAATTGGCTGGTGGTCTTCTCTATCTTCTCTCTG 186
Db 1125 MMYKMGGKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMMKKKMM 1066
Qy 187 GCTTTACTGTCTGTCTAACCATTCCAGCTTCGCTTACGAGGTGGCGAACGTTGCCGGG 246
Db 1065 NAXAKDKKKKKKAKAKMM 1010
Qy 247 ATGTACCATGTACGAAACGACTGCTCAAACCTCAAGCATTTGTGTATGAGGCGGACATG 306
Db 1009 KBKKMMCMMMKRCBMNKKKKKKKCNKCNKCNKCNKCNKCNKCNKCNKCNKCNKCNK 950
Qy 307 ATCATGCACACCCCGGGTGGT-GCCCTGCGTTTCGGGAGAACAACTCTTCCCGTCTCTG 365
Db 949 KKKMMMMAMCKYMBMCMMMMMMMMAKKKKKTKMTCKMKMKMKMKMKMKKKKKKKKK 890
Qy 366 GGTAGCGCTCACCCCGGCTCGCGGCTAGGAAACGCGCATCCCCACTACAACAATACG 425
Db 889 GKKKKCNKMKKKMKMKKCGMTKGTMMMGKMGCMMGKTTTGTMTTMMGMRMGGGAT 830
Qy 426 AGCCACGTCGATTTGCTGCTTGGGGCGGCTGCTTTC 462
Db 829 MKGTKKKKMAKMTWGMCMGMMGGGGGMMTKTKTKTTTTM 793

Search completed: December 20, 2003, 06:54:45
Job time : 1638.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.8681 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-13

Perfect score: 636

Sequence: 1 ATGCTGGTAAAGCCATCGA.....TGATGACTGGTACTAATAG 636

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCUTS_COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfilees1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	636	3	US-08-612-973-13
2	636	100.0	636	3	US-08-927-597-13
3	612.8	96.4	633	3	US-08-612-973-7
4	612.8	96.4	633	3	US-08-927-597-7
5	604	95.0	795	3	US-08-612-973-5
6	604	95.0	795	3	US-08-927-597-5
7	601.6	94.6	2082	3	US-08-612-973-47
8	601.6	94.6	2082	3	US-08-927-597-47
9	601.6	94.6	2433	3	US-08-612-973-49
10	601.6	94.6	2433	3	US-08-927-597-49
11	553.6	87.0	1539	2	US-08-470-426B-17
12	553.6	87.0	1863	2	US-08-470-426B-14
13	550.4	86.5	932	1	US-08-081-072-15
14	550.4	86.5	932	1	US-08-449-093A-15
15	548.8	86.3	1037	1	US-08-462-195-1
16	548.8	86.3	1037	3	US-08-636-883-1
17	548.8	86.3	1037	3	US-09-127-829-1
18	548.8	86.3	2116	3	US-08-191-160-21
19	545.6	85.8	9595	3	US-09-014-416-4
20	545.6	85.8	9599	3	US-09-014-416-6
21	534.4	84.0	1167	1	US-08-324-977-9
22	534.4	84.0	1167	2	US-08-384-616-9
23	534.4	84.0	1167	2	US-08-904-686A-9
24	534.4	84.0	1167	3	US-09-315-850-9
25	534.4	84.0	1499	1	US-08-324-977-3
26	534.4	84.0	1499	2	US-08-616-3
27	534.4	84.0	1499	2	US-08-904-686A-3

28	534.4	84.0	1499	3	US-09-315-850-3	Sequence 3, Appli
29	534.4	84.0	6039	1	US-08-324-977-11	Sequence 11, Appl
30	534.4	84.0	6039	1	US-08-384-616-11	Sequence 11, Appl
31	534.4	84.0	6039	2	US-08-904-686A-11	Sequence 11, Appl
32	534.4	84.0	6039	3	US-09-315-850-11	Sequence 11, Appl
33	534.4	84.0	9030	1	US-08-324-977-13	Sequence 13, Appl
34	534.4	84.0	9030	2	US-08-384-616-13	Sequence 13, Appl
35	534.4	84.0	9030	2	US-08-904-686A-13	Sequence 13, Appl
36	534.4	84.0	9030	3	US-09-315-850-13	Sequence 13, Appl
37	534.4	84.0	9416	1	US-08-324-977-1	Sequence 1, Appli
38	534.4	84.0	9416	2	US-08-384-616-1	Sequence 1, Appli
39	534.4	84.0	9416	2	US-08-904-686A-1	Sequence 1, Appli
40	534.4	84.0	9416	3	US-09-315-850-1	Sequence 1, Appli
41	534.4	84.0	9416	4	US-08-823-895A-27	Sequence 27, Appl
42	534.4	84.0	9472	4	US-08-150-204E-96	Sequence 96, Appl
43	532.8	83.8	1037	1	US-08-462-195-3	Sequence 3, Appli
44	532.8	83.8	1037	2	US-08-636-883-3	Sequence 3, Appli
45	532.8	83.8	1037	3	US-09-127-829-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-612-973-13
; Sequence 13, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, RONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630

Mon Dec 22 13:28:35 2003

```
US-08-612-973-13
Query Match      100.0%; Score 636; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 3.3e-164;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAGGCCATCATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGCTGGGTAAGGCCATCATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCCCCCTTAGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGGGTT 120
DB 61 CCGCTCGTGGGCCCCCTTAGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGGGTT 120
QY 121 CTGGAAGACGGGTGAACATGATCAAGGAAATTCCTTCTCTATCTTC 180
DB 121 CTGGAAGACGGGTGAACATGATCAAGGAAATTCCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTACTGCTCTCTCTAACCATTCCAGCTTCGGCTTACGAGTGCACACGTG 240
DB 181 CTCTTGGCTTTACTGCTCTCTCTAACCATTCCAGCTTCGGCTTACGAGTGCACACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATAGGACGG 300
DB 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATAGGACGG 300
QY 301 GACATGATCATGACACCCCGGTCGCTCCCTCGGTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACCCCGGTCGCTCCCTCGGTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCGGCTAGGAAACGCCAGCATCCCACTACAACA 420
DB 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCGGCTAGGAAACGCCAGCATCCCACTACAACA 420
QY 421 ATACGAGCCCAACGCTCGAATTTGCTGGTTGGGGGGGCTCTTCTTCTCGCTATGTAGTG 480
DB 421 ATACGAGCCCAACGCTCGAATTTGCTGGTTGGGGGGGCTCTTCTTCTCGCTATGTAGTG 480
QY 481 GGGGATCTCTCGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 481 GGGGATCTCTCGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/612,973
APPLICATION NUMBER:
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..630
US-08-927-597-13
Query Match      100.0%; Score 636; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 3.3e-164;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAGGCCATCATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGCTGGGTAAGGCCATCATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCCCCCTTAGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGGGTT 120
DB 61 CCGCTCGTGGGCCCCCTTAGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGGGTT 120
QY 121 CTGGAAGACGGGTGAACATGATCAAGGAAATTCCTTCTCTATCTTC 180
DB 121 CTGGAAGACGGGTGAACATGATCAAGGAAATTCCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTACTGCTCTCTCTAACCATTCCAGCTTCGGCTTACGAGTGCACACGTG 240
DB 181 CTCTTGGCTTTACTGCTCTCTCTAACCATTCCAGCTTCGGCTTACGAGTGCACACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATAGGACGG 300
DB 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATAGGACGG 300
QY 301 GACATGATCATGACACCCCGGTCGCTCCCTCGGTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACCCCGGTCGCTCCCTCGGTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCGGCTAGGAAACGCCAGCATCCCACTACAACA 420
DB 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCGGCTAGGAAACGCCAGCATCCCACTACAACA 420
QY 421 ATACGAGCCCAACGCTCGAATTTGCTGGTTGGGGGGGCTCTTCTTCTCGCTATGTAGTG 480
DB 421 ATACGAGCCCAACGCTCGAATTTGCTGGTTGGGGGGGCTCTTCTTCTCGCTATGTAGTG 480
QY 481 GGGGATCTCTCGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 481 GGGGATCTCTCGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

RESULT 2
US-08-927-597-13
; Sequence 13, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


QY 541 CATGACGGTGCAGGACTGCAATTCGTCATCTATCCCGGCCACATAACGGGTACCGGT 600
Db 541 CATGACGGTGCAGGACTGCAATTCGTCATCTATCCCGGCCACATAACGGGTACCGGT 600
QY 601 ATGGCTTGGGATATGATGAACTGGTACTAATAG 636
Db 601 ATGGCTTGGGATATGATGAACTGGTACTAATAG 636

RESULT 3
US-08-612-973-7
; Sequence 7, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..627
US-08-612-973-7

Query Match 96.4%; Score 612.8; DB 3; Length 633;
Best Local Similarity 98.1%; Pred. No. 7e-158;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATGCTGGTAAAGCATCGATACCTTTACGTGGGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTCATCGATACCTTTACGTGGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCCTCGTCCGGCGCCCTAGGGGCGCTGCCAGGCGCTGGCGCATGCGTCCGGGTT 120
Db 61 CCCTCGTCCGGCGCCCTAGGGGCGCTGCCAGGCGCTGGCGCATGCGTCCGGGTT 120

QY 121 CTGGAAGAGCGCGCTGAACATATGCAACAGGGAATTTCCCTGGTGGCTCTTTCTCTATCTTC 180
Db 121 CTGGAAGAGCGCGCTGAACATATGCAACAGGGAATTTCCCTGGTGGCTCTTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTTACTGTCTCTTAACCATTTCCAGCTTTCCGCTTACGAGGTGCCCAACGTG 240
Db 181 CTCTTGGCTTTTACTGTCTCTTAACCATTTCCAGCTTTCCGCTTACGAGGTGCCCAACGTG 240
QY 241 TCCGGGATGTACCATGTCTCAGAACGATCTCTCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGGATGTACCATGTCTCAGAACGATCTCTCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATCATGACACACCCCGGCTCGCTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
Db 301 GACATCATGACACACCCCGGCTCGCTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGCTCGGGCTAGGAACGCGAGCATCCCCACTACAACA 420
Db 361 TGCTGGGTAGCGCTCACCCCGCTCGGGCTAGGAACGCGAGCATCCCCACTACAACA 420
QY 421 ATACGACGCCACGTTCGATTTGCTTCGTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGTTCGATTTGCTTCGTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG 480
QY 481 GGGGATCTCTGGGATCTGTCTTCCTCGTCTCCAGCTGTTACCAATCTCGCCCTCGCGG 540
Db 481 GGGGATCTCTGGGATCTGTCTTCCTCGTCTCCAGCTGTTACCAATCTCGCCCTCGCGG 540
QY 541 CATGACACGGTGCAGGACTGCAATTCCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
Db 541 CATGACACGGTGCAGGACTGCAATTCCTCAATCTATCCCGGCCACATAACGGGTACCGT 600
QY 601 ATGGCTTGGGATATGATGAACTGGTACTA 632
Db 601 ATGGCTTGGGATATGATGAACTGGTACTA 632

RESULT 4
US-08-927-597-7
; Sequence 7, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:

Mon Dec 22 13:28:35 2003

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 633 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..630
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..627
US-08-927-597-7

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
US-08-612-973-5

Query Match 96.4%; Score 612.8; DB 3; Length 633;
Best Local Similarity 98.1%; Pred. No. 7e-158;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAGCCATCGATACCTTTAAGTGGGCTTGGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTCATCGATACCTTTAAGTGGGCTTGGCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGGCGGCGGCTAGGGGCGCTGCCAGGCGCTGGCGACCTGGCGTCCGGGTT 120
DB 61 CCGCTCGTGGGCGGCGGCTAGGGGCGCTGCCAGGCGCTGGCGACCTGGCGTCCGGGTT 120

QY 121 CTGGAAGACGCGGTGAACATATGCAACAGGGAATTTGCTGCTTTCTATCTTC 180
DB 121 CTGGAAGACGCGGTGAACATATGCAACAGGGAATTTGCTGCTTTCTATCTTC 180

QY 181 CTCTGGCTTTTACTGCTGCTTAAACATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
DB 181 CTCTGGCTTTTACTGCTGCTTAAACATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240

QY 241 TCCGGGATGTACCATGTGTCAGGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTGTCAGGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

QY 301 GACATGATCATGCAACCCCGGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCAACCCCGGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCGGCTAGGAACGCGACGATCCCACTACAA 420
DB 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCGGCTAGGAACGCGACGATCCCACTACAA 420

QY 421 ATACGACGCGCGTGGATTTGCTGCTGGGGCGGCTGCTTTCTGCTGGCTATGATGTTG 480
DB 421 ATACGACGCGCGTGGATTTGCTGCTGGGGCGGCTGCTTTCTGCTGGCTATGATGTTG 480

QY 481 GGGGATCTCTGGGATCTGCTTCTCTCTCCAGCTGTTCCACATCTCGCTCGCGG 540
DB 481 GGGGATCTCTGGGATCTGCTTCTCTCTCCAGCTGTTCCACATCTCGCTCGCGG 540

QY 541 CATGAGACGCGTGGAGACTGCAATTTGCTCAATCTATCCCGGCAATTAACGGGTACCGT 600
DB 541 CATGAGACGCGTGGAGACTGCAATTTGCTCAATCTATCCCGGCAATTAACGGGTACCGT 600

QY 601 ATGGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632
DB 601 ATGGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632

RESULT 5
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGCTGGGTAAGCCATCGATACCTTTAAGTGGGCTTGGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTCATCGATACCTTTAAGTGGGCTTGGCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGGCGGCGGCTAGGGGCGCTGCCAGGCGCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCGCTCGTGGGCGGCGGCTAGGGGCGCTGCCAGGCGCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGGAAGACGCGGTGAACATATGCAACAGGGAATTTGCTGCTTTCTATCTTC 180
DB 121 CTGGAAGACGCGGTGAACATATGCAACAGGGAATTTGCTGCTTTCTATCTTC 180

QY 181 CTCTGGCTTTTACTGCTGCTTAAACATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240
DB 181 CTCTGGCTTTTACTGCTGCTTAAACATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTTG 240

QY 241 TCCGGGATGTACCATGTGTCAGGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTGTCAGGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

QY 301 GACATGATCATGCAACCCCGGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCAACCCCGGGTGGTCCCTCGGTTCCGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGCACACCCCGGGTGGTGGTCCCTCGCTTGGGAGAACAACTCTTCCCGC 360
Qy 361 TGTGGGTAGCGCTACCCCGCCAGCTCGCGGCTAGGAACGCCAGCATFCCCACTACAACA 420
Db 361 TGTGGGTAGCGCTACCCCGCCAGCTCGCGGCTAGGAACGCCAGCGTCCCGCCAGGACA 420
Qy 421 ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCGTTCCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCGTTCCCGCTATGTACGTG 480
Qy 481 GGGGATCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Db 481 GGGGATCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Qy 541 CATGACAGCGTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600
Db 541 CATGACAGCGTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600
Qy 601 ATGGCTTGGGATATGATGAACCTGTT 628
Db 601 ATGGCTTGGGATATGATGAACCTGTT 628

RESULT 6

US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792

; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789
; US-08-927-597-5

Query Match 95.0%; Score 604; DB 3; Length 795;
Best Local Similarity 97.6%; Pred. No. 1.9e-155;
Matches 613; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 ATGCTGGGTAAAGCCATCGATACCTTACGTTGGCGGCTTGGCCGACCTCGTGGGGTACATT 60
Db 1 ATGCTGGGTAAAGTCAATCGATACCTTACATGCGGCTTGGCCGACCTCGTGGGGTACATT 60
Qy 61 CCCTCGTGGCGCCCGCTAGGGGGCGGTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
Db 61 CCCTCGTGGCGCCCGCTAGGGGGCGGTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
Qy 121 CTGGAAGAGCGGCGTGAACATGATGCAACAGGGAATTTGCGTGGTTCCTTCTATCTTC 180
Db 121 CTGGAAGAGCGGCGTGAACATGATGCAACAGGGAATTTGCGTGGTTCCTTCTATCTTC 180
Qy 181 CTCTGGCTTTACTGTCCTGTCTAAACATTTCCAGCTTCCGCTTACGAGGTGCGCAACGTG 240
Db 181 CTCTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTCCGCAACGTG 240
Qy 241 TCCGGATGTACCATGTCTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGATGTACCATGTCTACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Qy 301 GACATGATCATGCACACCCCGGGTGGCGTCCCTGGGTTCCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGCGTCCCTGGGTTCCGGAGAACAACTCTTCCCGC 360
Qy 361 TGTGGGTAGCGCTCACCCCGCCAGCTCGCGGCTAGGAACGCCAGCATCCCCACTACAACA 420
Db 361 TGTGGGTAGCGCTCACCCCGCCAGCTCGCGGCTAGGAACGCCAGCATCCCCACTACAACA 420
Qy 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
Qy 481 GGGGATCTCTGGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Db 481 GGGGATCTCTGGGATCTGCTTCTCTCGTCTCCAGCTGTTCAACCATCTCGCCTCGCGG 540
Qy 541 CATGACAGCGTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600
Db 541 CATGACAGCGTGACGACTGCAATTTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600
Qy 601 ATGGCTTGGGATATGATGAACCTGTT 628
Db 601 ATGGCTTGGGATATGATGAACCTGTT 628

RESULT 7

US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
NAME/KEY: mat peptide
LOCATION: 1..2076
US-08-612-973-47

Query Match 94.6%; Score 601.6; DB 3; Length 2082;
Best Local Similarity 97.8%; Pred. No. 1.1e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5 TGGGTAAGCCATCGATACCCCTTACGTGCGGCTTCGCGGACCTCGTGGGTACATTCGCG 64
DB 5 TGGGTAAGCTCATGATACCCCTTACATGCGGCTTCGCGGACCTCGTGGGTACATTCGCG 64
QY 65 TGGTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCGGGTCTG 124
DB 65 TGGTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCGGGTCTG 124
QY 125 AAGACGGCGTGAACATGATCAACAGGGAATTTGCCGTGGTGTCTTTCTTAATCTTCT 184
DB 125 AAGACGGCGTGAACATGATCAACAGGGAATTTGCCGTGGTGTCTTTCTTAATCTTCT 184
QY 185 TGGCTTTACTGTCTGTCTAACCATTCAGCTTCGGCTTACGAGTGCGCAACGTGTCGG 244
DB 185 TGGCTTTACTGTCTGTCTAACCATTCAGCTTCGGCTTACGAGTGCGCAACGTGTCGG 244
QY 245 GGATGTACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCGGACA 304
DB 245 GGATGTACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCGGACA 304
QY 305 TGATCATGACACCCCGGGTGGTCCCTGCTTGGGAGAACACTTCTCCCGCTGCT 364
DB 305 TGATCATGACACCCCGGGTGGTCCCTGCTTGGGAGAACACTTCTCCCGCTGCT 364
QY 365 GGGTAGGCTTACCCCGGCTGCTGCTTGGGAGAACACTTCTCCCGCTGCT 424
DB 365 GGGTAGGCTTACCCCGGCTGCTGCTTGGGAGAACACTTCTCCCGCTGCT 424
QY 425 GAGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTTGGGG 484
DB 425 GAGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTTGGGG 484
QY 485 ATCTCTGGGATGTCCTTCTCTGCTTCCAGCTGTTTCAACCATCTCGCCTCGCGGATG 544
DB 485 ACCTCTGGGATGTCCTTCTCTGCTTCCAGCTGTTTCAACCATCTCGCCTCGCGGATG 544

545 AGACGGTGCGAGGACTGCAATTCCTCAATCTATCCGCGCCACATAACGGGTACCGGTATGG 604
545 AGACGGTGCGAGGACTGCAATTCCTCAATCTATCCGCGCCACATAACGGGTACCGGTATGG 604
605 CTTGGGATATGATGATGAAGTGGT 628
605 CTTGGGATATGATGATGAAGTGGT 628

RESULT 8
US-08-927-597-47
Sequence 47, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
NAME/KEY: mat peptide
LOCATION: 1..2076
US-08-927-597-47

Query Match 94.6%; Score 601.6; DB 3; Length 2082;
Best Local Similarity 97.8%; Pred. No. 1.1e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5 TGGGTAAGCCATCGATACCCCTTACGTGCGGCTTCGCGGACCTCGTGGGTACATTCGCG 64
DB 5 TGGGTAAGCTCATGATACCCCTTACATGCGGCTTCGCGGACCTCGTGGGTACATTCGCG 64
QY 65 TGGTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCGGGTCTG 124
DB 65 TGGTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGCTCGGGTCTG 124
QY 125 AAGACGGCGTGAACATGATCAACAGGGAATTTGCCGTGGTGTCTTTCTTAATCTTCT 184
DB 125 AAGACGGCGTGAACATGATCAACAGGGAATTTGCCGTGGTGTCTTTCTTAATCTTCT 184
QY 185 TGGCTTTACTGTCTGTCTAACCATTCAGCTTCGGCTTACGAGTGCGCAACGTGTCGG 244
DB 185 TGGCTTTACTGTCTGTCTAACCATTCAGCTTCGGCTTACGAGTGCGCAACGTGTCGG 244
QY 245 GGATGTACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCGGACA 304
DB 245 GGATGTACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGAGCGGACA 304
QY 305 TGATCATGACACCCCGGGTGGTCCCTGCTTGGGAGAACACTTCTCCCGCTGCT 364
DB 305 TGATCATGACACCCCGGGTGGTCCCTGCTTGGGAGAACACTTCTCCCGCTGCT 364
QY 365 GGGTAGGCTTACCCCGGCTGCTGCTTGGGAGAACACTTCTCCCGCTGCT 424
DB 365 GGGTAGGCTTACCCCGGCTGCTGCTTGGGAGAACACTTCTCCCGCTGCT 424
QY 425 GAGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTTGGGG 484
DB 425 GAGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTTGGGG 484
QY 485 ATCTCTGGGATGTCCTTCTCTGCTTCCAGCTGTTTCAACCATCTCGCCTCGCGGATG 544
DB 485 ACCTCTGGGATGTCCTTCTCTGCTTCCAGCTGTTTCAACCATCTCGCCTCGCGGATG 544
```

Db	65		TGTTGGGCGCCCTTAGGGGGCGTGCAGGGCCCTTGGCGCATGGCGTCCGGTTCGTG	124
Qy	125		AAGCGGCGTGAATATATGAACAGGGAATTTGCCGTGTCTTTCTCTATCTTCTCT	184
Db	125		AGGACGGCGTGAACATATGAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTCTCT	184
Qy	185		TGGCTTTTACTGTCCTGTCTTAACCATTCACGACTTCGCGTTACGAGGTGCGCAACGTGTCCG	244
Db	185		TGGCTTTTGTGTCCTGTCTGACCGTTTCAGCTTTCGCGTTATGAAGTGCACGCTGTCCG	244
Qy	245		GGATGTACCATGTCAAGAACGACTCTCCAACTCAAGCATTTGTATAGGACGCGGACA	304
Db	245		GGATGTATCCATGTACAGAAAGACTGCTCCAACTCAAGCATTTGTATAGGACGCGGACA	304
Qy	305		TGATCATGCACACCCCGGGTGCCTCGGTTCCGGAGAACAACTCTTCCGCTGCT	364
Db	305		TGATCATGCACACCCCGGGTGCCTCGGTTCCGGAGAACAACTCTTCCGCTGCT	364
Qy	365		GGGTAGCGCTCACCCCCACGCTCGGGGTAGGAAAGCCAGCATCCCACTACAACAATAC	424
Db	365		GGGTAGCGCTCACCCCCACGCTCGGAGTAGAAGCAGCGTCCCAACCAACAATAC	424
Qy	425		GACGCCAGTCGATTTGCTCGTTGGGGCGGTGCTTTCTGTTCGCTATATAGTGGGG	484
Db	425		GACGCCAGTCGATTTGCTCGTTGGGGCGGTGCTTTCTGTTCGCTATATGTCGGGG	484
Qy	485		ATCTCTGGGATCTGCTTCTCTGCTCTCCAGCTGTTTCAACCATCTCGCCTCGCGGGATG	544
Db	485		ACTCTGGGGATCTGCTTCTCTCGTCTCCAGCTGTTTCAACCATCTCGCCTCGCGGCATG	544
Qy	545		AGACGGTGCAGACTGCAAATTTGCTCAATCTATCCGGCCACATAAGCGGTACCGTATGG	604
Db	545		AGACGGTGCAGACTGCAAATTTGCTCAATCTATCCGGCCACATAAGCGGTACCGTATGG	604
Qy	605		CTTTGGGATATGATGAAGTGGT	628
Db	605		CTTTGGGATATGATGAAGTGGT	628

RESULT 9

US-08-612-973-49
Sequence 49, Application US/08612973
Patent No. 6150134
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-927-597-49

Query Match 94.6%; Score 601.6; DB 3; Length 2433;
Best Local Similarity 97.8%; Pred. No. 1.2e-154;
Matches 610; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 TGGGTAAAGCCATCGATACCTTACGTGGCGCTTCGCCGACCTCGTGGGGTACATTCCGC 64
DB 356 TGGGTAAAGTCATCGATACCTTACGTGGCGCTTCGCCGACCTCGTGGGGTACATTCCGC 415
QY 65 TCCTCGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 124
DB 416 TCCTCGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTGG 475
QY 125 AAGACGGCGTGAATCATGCAAGGGAATTCGCTGTTGCTCTTCTATCTTCTCTCT 184
DB 476 AGGACGGCGTGAATCATGCAAGGGAATTCGCGGTTGCTCTTCTATCTTCTCTCT 535
QY 185 TGGCTTTACTGCTCTGCTTAACCATTCAGCTTCCGCTTACGAGGTGCGCAACGTTGTCG 244
DB 536 TGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACACGTTGTCG 595
QY 245 GGATGTACCATGTGCAAGACGCTGCTCAACCTCAAGCATTTGTATGAGGCGCGGACA 304
DB 596 GGATGTACCATGTGCAAGACGCTGCTCAACCTCAAGCATTTGTATGAGGCGCGGACA 655

QY 305 TGATCATGCACACCCCGGGTGGTGGCTTTCGGGAGAAACAATCTTCCCGTCTGT 364
DB 656 TGATCATGCACACCCCGGGTGGTGGCTTTCGGGAGAAACAATCTTCCCGTCTGT 715
QY 365 GGGTAGCGGTCAACCCCGACGCTCGCGGTAGGAAGCCAGCATTCGCCATACAAATAC 424
DB 716 GGGTAGCGGTCAACCCCGACGCTCGCGGTAGGAAGCCAGCATTCGCCATACAAATAC 775
QY 425 GAGCCACGTCGATTTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 484
DB 776 GAGCCACGTCGATTTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 835
QY 485 ATCTCTCGGATCTGCTTCTCTCCAGCTTCTCCAGCTGTTTCCACCATTCGCTCGCGCATG 544
DB 836 ACCTCTCGGATCTGCTTCTCTCCAGCTTCTCCAGCTGTTTCCACCATTCGCTCGCGCATG 895
QY 545 AGACGGTGCAGGACTGCAATTCGCTATTCGCCGCGCACATAACGGGTCAACGTTATGG 604
DB 896 AGACGGTGCAGGACTGCAATTCGCTATTCGCCGCGCACATAACGGGTCAACGTTATGG 955
QY 605 CTTGGGATATGATGATGAATGTT 628
DB 956 CTTGGGATATGATGATGAATGTT 979
RESULT 11
US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beverage, DeGrandi, Weilacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-470-426B-17

Query Match 87.0%; Score 553.6; DB 2; Length 1539;
Best Local Similarity 92.9%; Pred. No. 1.2e-141;
Matches 580; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 5 TGGGTAAGGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATTCGCCG 64
DB TGGGTAAGGTCATGATACCCCTTACATCGGCTTCGCCGATCTCATGGGGTATATCCCC 415
QY 65 TCGTGGGCGCCCTTAGGGGGCTGCGAGGCGCTGGGCGATGGCGTCCGGGTTCTGG 124
DB TCGTGGGCGCCCTTAGGGGGCTGCGAGGCGCTGGGCGATGGCGTCCGGGTTCTGG 475
QY 125 AAGACGGCTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCT 184
DB AGACGGCGTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCTCT 535
QY 185 TGGCTTTACTGCTCTCTAAACATTCAGCTTCGCTTACGAGGTGCGCAAGTGTCCG 244
DB TGGCTTTGCTGCTCTTTGACCATCCAGCTTCGCTTATGAAGTGGCAAGTGTCCG 595
QY 245 GGATGTACATGTACGAAAGCTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 304
DB GGATATACCATGTACGAAAGCTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 655
QY 305 TGATCATGCACACCCCGGCGTGGCTGCTGCGGTAGGAACGCCAGCATCCCACTACAA 364
DB TGATCATGCATCTCCCGGCGTGGCTGCTGCGGTAGGAACGCCAGCATCCCACTACAA 715
QY 365 GGGTAGCGCTACCCCGGCTAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 424
DB GGGTAGCGCTACCCCGGCTAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 775
QY 425 GACGCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
DB GACGCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
QY 485 ATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
DB ATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAACGGGTACCGTATG 604
DB AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAATTCAGTACCGGATG 955
QY 605 CTTGGGATATGATGAAGTGT 628
DB CTTGGGATATGATGAAGTGT 979

RESULT 12

US-08-470-426B-14
Sequence 14, Application US/08470426B
Patent No. 5856458
GENERAL INFORMATION:
APPLICANT: Okamoto, Hiroaki
APPLICANT: Nakamura, Tetsuo
TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
TITLE OF INVENTION: HEPATITIS VIRUS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
ADDRESSEE: L.L.P.
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,426B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-153402
FILING DATE: 12-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Weilacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/59-47083.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-470-426B-14

Query Match 87.0%; Score 553.6; DB 2; Length 1863;
Best Local Similarity 92.9%; Pred. No. 1.3e-141;
Matches 580; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 5 TGGGTAAGGCCATCGATACCCCTTACGTGGGCTTCGCCGACCTCGTGGGTACATTCGCCG 64
DB TGGGTAAGGTCATGATACCCCTTACATCGGCTTCGCCGATCTCATGGGGTATATCCCC 739
QY 65 TCGTGGGCGCCCTTAGGGGGCTGCGAGGCGCTGGGCGATGGCGTCCGGGTTCTGG 124
DB TCGTGGGCGCCCTTAGGGGGCTGCGAGGCGCTGGGCGATGGCGTCCGGGTTCTGG 799
QY 125 AAGACGGCGTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCT 184
DB AGACGGCGTGAACATGCAACAGGGAATTCGCTGTTGCTTCTCTATCTCTCTCTCT 859
QY 185 TGGCTTTACTGCTCTCTAAACATTCAGCTTCGCTTACGAGGTGCGCAAGTGTCCG 244
DB TGGCTTTGCTGCTCTTTGACCATCCAGCTTCGCTTATGAAGTGGCAAGTGTCCG 919
QY 245 GGATGTACATGTACGAAAGCTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 304
DB GGATATACCATGTACGAAAGCTGCTCCAACTCAAGCATTTGTTATGAGGAGCGGACA 979
QY 305 TGATCATGCACACCCCGGCGTGGCTGCTGCGGTAGGAACGCCAGCATCCCACTACAA 364
DB TGATCATGCATCTCCCGGCGTGGCTGCTGCGGTAGGAACGCCAGCATCCCACTACAA 1039
QY 365 GGGTAGCGCTACCCCGGCTAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 424
DB GGGTAGCGCTACCCCGGCTAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1099
QY 425 GACGCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
DB GACGCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
QY 485 ATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
DB ATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
QY 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAACGGGTACCGTATG 604
DB AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAATTCAGTACCGGATG 1279
QY 605 CTTGGGATATGATGAAGTGT 628
DB CTTGGGATATGATGAAGTGT 1303

RESULT 13

```
US-08-081-072-15
; Sequence 15, Application US/08081072
; Patent No. 5641654
; GENERAL INFORMATION:
; APPLICANT: No. 5641654oru MAKI, Kenjiro YAMAGUCHI, Ayumi
; APPLICANT: TOYOSHIMA, and Michinori KOHARA
; TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC
; TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS
; TITLE OF INVENTION: DIAGNOSIS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, Bronstein, Roberts & Cushman
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.50inch, 1.4Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh Classic
; SOFTWARE: Microsoft Word Version 4.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/081,072
; FILING DATE: June 22, 1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 932 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; US-08-081-072-15

Query Match      86.5%; Score 550.4; DB 1; Length 932;
Best Local Similarity 92.6%; Pred. No. 7.9e-141;
Matches 578; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY      5 TGGGTAAGGCATCGATACCCCTTACGTCGGCTTTCGCGACCTCGTGGGTACATTCGCG 64
DB      10 TGGGTAAGGTATCATGATACCCCTCACATGCGGCTTCGCCGACCTCATGCGGTACATTCGCG 69

QY      65 TCGTCGGCGCCCTAGGGCGCTGCCAGGCGCTTCCGACATTCGTAAGGATTCGTAAGGCGGATTCGCG 124
DB      70 TTGTTCGGCGCCCTAGGGCGCTGCCAGGCGCTTCCGACATTCGTAAGGATTCGTAAGGCGGATTCGCG 129

QY      125 AAGACGGCGTGAATTCGTAAGGCGCTTCCGACATTCGTAAGGCGGATTCGTAAGGCGGATTCGCG 184
DB      130 AAGACGGCGTGAATTCGTAAGGCGCTTCCGACATTCGTAAGGCGGATTCGTAAGGCGGATTCGCG 189

QY      185 TGGCTTTACTGTCCTGTCTTAACTTCCAGTTCGCTTACAGTTCGCTTACAGTTCGCTTACAGTTCGCTT 244
DB      190 TGGCTTTACTGTCCTGTCTTAACTTCCAGTTCGCTTACAGTTCGCTTACAGTTCGCTTACAGTTCGCTT 249

QY      245 GATGTACCATGTACGACGATGCTCCAACTCAAGCATTCGTAAGGATTCGTAAGGCGGATTCGCG 304
DB      250 GATGTACCATGTACGACGATGCTCCAACTCAAGCATTCGTAAGGATTCGTAAGGCGGATTCGCG 309

QY      305 TGATCATGACACCCCGGGTGGTCCCTTGGGTCGTAAGGCGGATTCGTAAGGCGGATTCGCG 364
DB      310 TGATCATGACACCCCGGGTGGTCCCTTGGGTCGTAAGGCGGATTCGTAAGGCGGATTCGCG 369

QY      365 GGGTAGGCGTACCCCGGCTCGCGGTAGGAGCCAGCATTCGCCACTACAAATAC 424
DB      370 GGGTAGGCGTACCCCGGCTCGCGGTAGGAGCCAGCATTCGCCACTACAAATAC 429

QY      425 GACGCCACGTCGATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGG 484
DB      430 GACGCCACGTCGATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGG 489

QY      485 ATCTCTGGGATCGTCTTCTCTGCTTCCAGCTGTTCCACATCTCGCTTCGCGCGCATG 544
```

```
DB      490 ATCTCTGTGATCTGTCTTCTCTCTCCAGCTGTTCACCTCTGTCGCGCATG 549
QY      545 AGACGGTGCAGGACTGCAATTCCTCAATCTATCCCGGCCACATACCGGTCACCGTATGG 604
DB      550 AGACAGTACAGGACTGCAATTCCTCAATCTATCCCGGCCACATTCGACAGGTATCGCATGG 609
QY      605 CTTGGGATATGATGATGAAGTGGT 628
DB      610 CTTGGGATATGATGATGAAGTGGT 633

RESULT 14
US-08-449-093A-15
; Sequence 15, Application US/08449093A
; Patent No. 5662906
; GENERAL INFORMATION:
; APPLICANT: No. 5662906oru MAKI, Kenjiro YAMAGUCHI, Ayumi
; APPLICANT: TOYOSHIMA, and Michinori KOHARA
; TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC
; TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS
; TITLE OF INVENTION: DIAGNOSIS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, Bronstein, Roberts & Cushman, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,093A
; FILING DATE: May 24, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/081,072
; FILING DATE: June 22, 1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/726,141
; FILING DATE: July 8, 1991
; CLASSIFICATION: 424
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 932 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; US-08-449-093A-15

Query Match      86.5%; Score 550.4; DB 1; Length 932;
Best Local Similarity 92.6%; Pred. No. 7.9e-141;
Matches 578; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY      5 TGGGTAAGGCATCGATACCCCTTACGTCGGCTTTCGCGACCTCGTGGGTACATTCGCG 64
DB      10 TGGGTAAGGTATCATGATACCCCTCACATGCGGCTTCGCCGACCTCATGCGGTACATTCGCG 69

QY      65 TCGTCGGCGCCCTAGGGCGCTGCCAGGCGCTTCCGACATTCGTAAGGATTCGTAAGGCGGATTCGCG 124
DB      70 TTGTTCGGCGCCCTAGGGCGCTGCCAGGCGCTTCCGACATTCGTAAGGATTCGTAAGGCGGATTCGCG 129

QY      125 AAGACGGCGTGAATTCGTAAGGCGCTTCCGACATTCGTAAGGCGGATTCGTAAGGCGGATTCGCG 184
DB      130 AAGACGGCGTGAATTCGTAAGGCGCTTCCGACATTCGTAAGGCGGATTCGTAAGGCGGATTCGCG 189

QY      185 TGGCTTTACTGTCCTGTCTTAACTTCCAGTTCGCTTACAGTTCGCTTACAGTTCGCTTACAGTTCGCTT 244
```


Db 190 TGGCTTTGCTGCTCTTTTACCATCCAGCTTCCGCTTATGAGTGGCGAAGTATCCG 249
Qy 245 GGATGTACCATGTACGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCGAGCGGACA 304
Db 250 GGATATACCATGTACGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCGAGCGGACA 309
Qy 305 TGATCATGCACACCCCGGGTGGTGGCTTCCGCTTCCGAGACAACTCTTCCCGTGTCT 364
Db 310 TGATCATGCATACCCCGGGTGGTGGCTTCCGCTTCCGAGACAACTCTTCCCGTGTCT 369
Qy 365 GGGTAGCGCTCACCCACGCTCGCGCTAGGAACCGCAGCATCCCACTACAAACAATAC 424
Db 370 GGCGACGCTCACTCCACGTTAGCGCGCAGGACACCGCTCCCACTACGACATAC 429
Qy 425 GACGCACGTCGATTTGCTGTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTGGGG 484
Db 430 GACGGCATGTGCAATTTGCTGTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTGGGG 489
Qy 485 ATCTCTGCGCATGTGCTTCTGCTTCCGCTTCCGAGCTTCCACCATCTCGCGCGCATG 544
Db 490 ATCTCTGCGCATGTGCTTCTGCTTCCGCTTCCGAGCTTCTTCTCACCTCGTCGCGCATG 549
Qy 545 AGACGGTGCAGGACTCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGG 604
Db 550 AGACAGTACAGGACTCGCACTGCTCAATCTATCCCGGCCACTTGACAGGTATCGCATGG 609
Qy 605 CTGGGATATGATGATGAAGTGT 628
Db 610 CTGGGATATGATGATGAAGTGT 633

RESULT 15

US-08-462-195-1
; Sequence 1, Application US/08462195
; Patent No. 5789544
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIHARU
; APPLICANT: HONDA, YOSHIKAZU
; APPLICANT: SEKI, MAKOTO
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,195
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE: 22-MAY-1995
; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5789544man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855.OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1037 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: pUC010
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..1036
; US-08-462-195-1

Query Match 86.3%; Score 548.8; DB 1; Length 1037;
Best Local Similarity 92.5%; Pred. No. 2.2e-140;
Matches 577; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
Qy 5 TGGGTAAGGCCATCGATACCCCTTACGTGGGGCTTTCGCCGACCTCGTGGGGGTACATTCGCG 64
Db 372 TGGGTAAGGTCTATCGATACCCCTTACATGCGGGCTTTCGCCGACCTCATGCGGGTACATCCCG 431
Qy 65 TGTGCGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGCGGATGCGGTCCGGGTCTGG 124
Db 432 TTGTGCGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGCGACATGTTGTGTCGGGGTCTGG 491
Qy 125 AAGACGGCGTGAATATGCAACAGGGAATTTGCTGTTGCTCTTCTCTATCTCTCTCT 184
Db 492 AGGACGGCGTGAATATGCAACAGGGAATTTGCGGGTTCCTTCTCTATCTCTCTCT 551
Qy 185 TGGCTTTACTGCTGCTGCTAACTTCAGCTTCGCTTACGAGTTCGCGGACGAGTGTCCG 244
Db 552 TAGCTCTGCTGCTGCTTTGACCATCCAGCTTCCGCTTATGAAGTGCAGCAAGTGTCCG 611
Qy 245 GGATGTACCATGTACGAACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGCGGACA 304
Db 612 GGATATACCATGTACGAACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGCGGAG 671
Qy 305 TGATCATGCACACCCCGGGTGGTGGCTTCCGAGTTCGAGAGAACAACTCTTCCCGTGTCT 364
Db 672 TGATCATGCATGCCCCCGGGTGGTGGCTTCCGAGTTCGAGAGAACAACTCTTCCCGTGTCT 731
Qy 365 GGGTAGCGCTCACCCCGCACGCTCGGGGTAGAAAGCCAGCATCCCACTACAAACAATAC 424
Db 732 GGGTAGCGCTCACTCCACGCTCGCGGCGAGGAATGCCAGCGTCCCACTACGACATTAC 791
Qy 425 GAGGCCACGTGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 484
Db 792 GAGGCCACGTGCACTTGTCTGTTGGAGCGGCTGCTTCTGCTCCGCTATGTAGTGGGG 851
Qy 485 ATCTCTGCGGATCTGCTTCTCTCTGCTCCAGCTGTTTCAACCATCTCGCTCGCGCGCATG 544
Db 852 ATCTCTGCGGATCTGTTTCTCTCTCTGCTCCAGCTGTTTCACTTCTGCTCCGCTCGCGCATG 911
Qy 545 AGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGG 604
Db 912 AGACAGTACAGGACTGCAACTGCTCAATCTATCCCGGCCACGTTATCAGGCCATCGTATGG 971
Qy 605 CTGGGATATGATGATGAAGTGT 628
Db 972 CTGGGATATGATGATGAAGTGT 995

Search completed: December 20, 2003, 07:03:03
Job time : 47.8681 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2737.48 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303a-21

Perfect score: 723

Sequence: 1 ATGTTGGTAGGTCATCGA.....TACTCTTGTCTCCCTAATAG 723

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	723	100.0	723	6	A48683	A48683 Sequence 21
2	723	100.0	723	6	AR157336	AR157336 Sequence
3	723	100.0	723	6	AX452770	AX452770 Sequence
4	723	100.0	723	6	AX685022	AX685022 Sequence
5	641	88.7	795	6	A48667	A48667 Sequence 5
6	641	88.7	795	6	AR157325	AR157325 Sequence
7	641	88.7	795	6	AX452754	AX452754 Sequence
8	641	88.7	795	6	AX685006	AX685006 Sequence
9	624.8	86.4	2082	6	A48709	A48709 Sequence 47
10	624.8	86.4	2082	6	AR157350	AR157350 Sequence
11	624.8	86.4	2082	6	AX452796	AX452796 Sequence
12	624.8	86.4	2082	6	AX685048	AX685048 Sequence
13	624.8	86.4	2433	6	A48711	A48711 Sequence 49
14	624.8	86.4	2433	6	AR157351	AR157351 Sequence
15	624.8	86.4	2433	6	AX452798	AX452798 Sequence
16	624.8	86.4	2433	6	AX685050	AX685050 Sequence
17	598.2	82.7	606	6	A48687	A48687 Sequence 25
18	598.2	82.7	606	6	AR157338	AR157338 Sequence
19	598.2	82.7	606	6	AX452774	AX452774 Sequence
20	598.2	82.7	606	6	AX685026	AX685026 Sequence
21	597	82.6	636	6	A48689	A48689 Sequence 27
22	597	82.6	636	6	AR157339	AR157339 Sequence
23	597	82.6	636	6	AX452776	AX452776 Sequence
24	597	82.6	636	6	AX685028	AX685028 Sequence
25	565.6	78.2	3296	14	AB008441	AB008441 Hepatitis
26	565.6	78.2	3296	14	AB008443	AB008443 Hepatitis
27	565	78.1	9418	14	HCV132996	HCV132996 Hepatitis
28	564	78.0	3296	14	AB008442	AB008442 Hepatitis
29	564	78.0	9379	14	AF207766	AF207766 Hepatitis
30	564	78.0	9379	14	HCVPOLYP	HCVPOLYP Hepatitis
31	562.4	77.8	9386	14	AF165056	AF165056 Hepatitis
32	560.8	77.6	3296	14	AB008446	AB008446 Hepatitis
33	560.8	77.6	9386	14	AF165055	AF165055 Hepatitis
34	559.2	77.3	3296	14	AB008447	AB008447 Hepatitis
35	559.2	77.3	9379	14	AF165052	AF165052 Hepatitis
36	559.2	77.3	9379	14	AF207761	AF207761 Hepatitis
37	559.2	77.3	9410	14	HPCKIR2	D50481 Hepatitis C
38	559.2	77.3	9605	6	AX739971	AX739971 Sequence
39	559.2	77.3	9605	14	HCV238799	HCV238799 Hepatitis
40	559.2	77.3	11076	6	AX036252	AX036252 Sequence
41	559.2	77.3	11076	6	AX036258	AX036258 Sequence
42	559.2	77.3	11076	6	AX036260	AX036260 Sequence
43	559.2	77.3	11076	6	AX036262	AX036262 Sequence
44	557.6	77.1	3296	14	AB008445	AB008445 Hepatitis
45	557.6	77.1	9033	14	HCV238800	HCV238800 Hepatitis

ALIGNMENTS

RESULT 1
A48683
LOCUS A48683 Sequence 21 from Patent WO9604385. 723 bp DNA linear PAT 07-MAR-1997
DEFINITION A48683
ACCESSION A48683
VERSION A48683.1 GI:2302396
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified.
REFERENCE 1 (bases 1 to 723)
Maertens,G., Bosman,F., De.M.G. and Buysse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 21 15-FEB-1996;

DB
801
661
QY
720

CAAGCGTCTCCACATCGTGGCGGGGCCCATTTGGGAGTCCTTGGCGGTCTCGCCTAC 66

```
Db 601 CAAGCTGTCTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTCGGGGTCTCGCCTAC 660
QY 661 TATTCCATGGTGGGAACTGGGCTAAGGTTTGTGATGCTACTCTTTTGTCTCCCTAA 720
Db 661 TATTCCATGGTGGGAACTGGGCTAAGGTTTGTGATGCTACTCTTTTGTCTCCCTAA 720
QY 721 TAG 723
Db 721 TAG 723

RESULT 3
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
INNOCENETICS N.V. (BE)
FEATURES
Location/Qualifiers
source 1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD38072.1"
/db_xref="GI:21712456"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAARALAHGVRVLEDD
VNYATGNLPGCSFIFLLALLSLTPASAYEVNRVSGMYHTVNDCSNSSIIVEAADM
IMHTGCVPCVRENSSRCWVALTPTLAARNASVPTTIRRHVDLSQFTISPRRHETV
ODNCNSIYPGHITGHRMADMMNWSPTTALVVSQLLRIPQAVDVNVAHGVLAGL
AYSVNGVWAKVLIIVLLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN

Query Match 100.0%; Score 723; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 7.4e-148;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGCGTTCGGGTT 120
Db 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCGTTCGGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGGTGTCTTTCTCTAFTTC 180
Db 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTCCCGGGTGTCTTTCTCTAFTTC 180
QY 181 CTCCTGGCTTCTGTCCTCTGACCGTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 240
Db 181 CTCCTGGCTTCTGTCCTCTGACCGTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 240
QY 241 TCCGGATGTACCATGTACAGAACACTGCTCCAACTCAAGCATTCGTATGAGGACGCG 300
Db 241 TCCGGATGTACCATGTACAGAACACTGCTCCAACTCAAGCATTCGTATGAGGACGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGTTCGGGAGAACAACTCTTCCCGC 360
```

```
Db 301 GACATGATCATGCACACCCCGGGTGGTGGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGTCTGGTAGCGCTCACCCCCACAGCTCGCAGCTAGGAAAGCCAGCGTCCCCACACGACA 420
Db 361 TGTCTGGTAGCGCTCACCCCCACAGCTCGCAGCTAGGAAAGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCGCAGCTCGATTCCAGCTGTTACCATCTCGCTCGCGGATGAGACGGTG 480
Db 421 ATACGACGCGCAGCTCGATTCCAGCTGTTACCATCTCGCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCGGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGGTTCGCTACACCGGCCCTGGTGTATCGCAGCTGCTCCGATCCCA 600
Db 541 ATGATGATGAACCTGGTTCGCTACACCGGCCCTGGTGTATCGCAGCTGCTCCGATCCCA 600
QY 601 CAAGCTGTCTGGACATGTTGGGGGGGCCAATTTGGGGAGTCTTGGCGGGTCTCGCCTAC 660
Db 601 CAAGCTGTCTGGACATGTTGGGGGGGCCAATTTGGGGAGTCTTGGCGGGTCTCGCCTAC 660
QY 661 TATTCCATGGTGGGAACTGGGCTAAGGTTTGTGATGCTACTCTTTTGTCTCCCTAA 720
Db 661 TATTCCATGGTGGGAACTGGGCTAAGGTTTGTGATGCTACTCTTTTGTCTCCCTAA 720
QY 721 TAG 723
Db 721 TAG 723

RESULT 4
AX685022
LOCUS AX685022 723 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 21 from Patent WO0205548.
ACCESSION AX685022
VERSION AX685022.1 GI:29371427
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; serNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Maertens,G., Bosman,F. and Buyse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 21 18-JUL-2002;
INNOCENETICS N.V. (BE)
FEATURES
Location/Qualifiers
source 1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD86525.1"
/db_xref="GI:29371428"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAARALAHGVRVLEDD
VNYATGNLPGCSFIFLLALLSLTPASAYEVNRVSGMYHTVNDCSNSSIIVEAADM
IMHTGCVPCVRENSSRCWVALTPTLAARNASVPTTIRRHVDLSQFTISPRRHETV
ODNCNSIYPGHITGHRMADMMNWSPTTALVVSQLLRIPQAVDVNVAHGVLAGL
AYSVNGVWAKVLIIVLLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN

Query Match 100.0%; Score 723; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 7.4e-148;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	ATGTTGGGTAAGGTCATCGATACCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1		
QY	61	CGGCTCGTCGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Db	61		
QY	121	CTGAGGACGCGGTGAATCATGCAACAGGGAATTTGCCCGGTGTCTTTCTCTATCTTC	180
Db	121		
QY	181	CTCTTGGCTTTGCTGCTCTGTCACCGTTCAGCTTCGGCTTATGAAGTGCGAACGCG	240
Db	181		
QY	241	TCCGGGATGTAACCATGTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG	300
Db	241		
QY	301	GACATGATCATGCACACCCCGGGTGGCTGGCTCGGTTCGGGAGAACAACTCTTCCCGC	360
Db	301		
QY	361	TGCTGGGTAGCGCTCACCCCGCGCTCGACGTAGGAAACGCCAGCGTCCCCACACGACA	420
Db	361		
QY	421	ATACGACGCCACGTGCAATCCAGCTGTTCAACATCTCGGCTCGCGGCGATCAGACGGTG	480
Db	421		
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTTGGGAT	540
Db	481		
QY	541	ATGATGATGAATGGTTCGCCTACAACGGCCCTGGTGTTATCGCAGCTGCTCCGGATCCCA	600
Db	541		
QY	601	CAAGCTGCTGTGACATGGTGGCGGGGCCATTGGGGAGTTCCTGGCGGGTCTCGCCTAC	660
Db	601		
QY	661	TATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGTTGATGCTACTCTTTGCTCCCTAA	720
Db	661		
QY	721	TAG	723
Db	721		
QY	721	TAG	723
Db	721		

RESULT 5	linear	DNA	795 bp	PAT 07-MAR-1997
A48667				
LOCUS				
DEFINITION	Sequence 5 from Patent WO9604385.			
ACCESSION	A48667			
VERSION	A48667.1			
KEYWORDS	GI:2302380			
SOURCE	unidentified			
ORGANISM	unidentified			
REFERENCE	1 (bases 1 to 795)			
AUTHORS	Maertens, G., Bosman, F., De, M.G. and Buyse, M.			
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE			
JOURNAL	Patent: WO 9604385-A 5 15-FEB-1996;			
COMMENT	INNOGENETICS NV (BE) Other publication CA 2172273 960215 Other publication AU 3382495 960304.			
FEATURES	Location/Qualifiers			
SOURCE	1..795			

```

/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

1..792
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03129.1"
/db_xref="GI:2302191"
/translation="MLGKFLDLTLCGFPADILGVYPLVCAPLGGRARALAHGVRVLEDDG
VNYATGNLPGGSFGIFLLALISCLTFSAYSEVRNVSGMYHVINDCSNSSIYVEADMM
IMHTPGCVPCVRENNSSRCWALPTLAARNASVETTTIRRHVDLLVGAAAFCSAMYV
GDLCSPLVLSOLFTISPRRHETVDCNSIYPGHITGRMAWDDMMNSPTTALIVVS
QLLRIPQAVDMVAGHWGLAGLAYISMVGNWAKVLIVMLLIFAP"

mat_peptide
1..798
/product="unnamed"

BASE COUNT      130 a      240 c      231 g      194 t
ORIGIN
Query Match      88.7%; Score 641; DB 6; Length 795;
Best Local Similarity 90.9%; Pred. No. 6,3e-130;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY      1  ATGTTGGGTAAAGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
DB      1  ATGTTGGGTAAAGTCAATCATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60

QY      61  CGGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGAGGCCCTGCGCGATGCGGCTCCGGGTT 120
DB      61  CGGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGAGGCCCTGCGCGATGCGGCTCCGGGTT 120

QY      121  CTGGAGACGGCTGAACATATGCAACAGGAAATTGCGGGTTGCTCTTTCTCTATCTTC 180
DB      121  CTGGAGACGGCTGAACATATGCAACAGGAAATTGCGGGTTGCTCTTTCTCTATCTTC 180

QY      181  CTCTTGCTTTGCTGCTCTGTCGACCGTTCCAGCTTCGCGTTATGAAGTGCACAAGCTG 240
DB      181  CTCTTGCTTTGCTGCTCTGTCGACCGTTCCAGCTTCGCGTTATGAAGTGCACAAGCTG 240

QY      241  TCCGGGATGTACCATGTACGAAACGACTGCTCCAACCTCAAGCAATGTGTATGAGCGACG 300
DB      241  TCCGGGATGTACCATGTACGAAACGACTGCTCCAACCTCAAGCAATGTGTATGAGCGACG 300

QY      301  GACATGATCATGCACACCCCGGGTGGTGCCTTCCGTTCCGGAGAACAACTCTTCCCGC 360
DB      301  GACATGATCATGCACACCCCGGGTGGTGCCTTCCGTTCCGGAGAACAACTCTTCCCGC 360

QY      361  TGCTGGGTAGCGCTCACCCACAGCTCGCAGCTAGGAACGCGCAGCGTCCCACACGACA 420
DB      361  TGCTGGGTAGCGCTCACCCACAGCTCGCAGCTAGGAACGCGCAGCGTCCCACACGACA 420

QY      421  ATACGACGCCACGTCGAT----- 438
DB      421  ATACGACGCCACGTCGATTTGCTTGGGGGGCGTGTCTTGTTCGCTATGTACGTG 480

QY      439  -----TCCCAAGCTGTTCAACCATCTCGCCTCCGCGG 468
DB      439  -----TCCCAAGCTGTTCAACCATCTCGCCTCCGCGG 468

QY      481  GGGGACCTCTGCGAATGTCTTCCTCGTCTCCAGCTGTTCAACCATCTCGCCTCCGCGG 540
DB      481  GGGGACCTCTGCGAATGTCTTCCTCGTCTCCAGCTGTTCAACCATCTCGCCTCCGCGG 540

QY      469  CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGGCACATACAGGGTCACCGT 528
DB      469  CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGGCACATACAGGGTCACCGT 528

QY      541  CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGGCACATACAGGGTCACCGT 600
DB      541  CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGGCACATACAGGGTCACCGT 600

QY      529  ATGGCTTGGGATATGATGATGAACCTGGTGCCTACAAACGGGCCCTGGTGTATCGCAGCTG 598
DB      529  ATGGCTTGGGATATGATGATGAACCTGGTGCCTACAAACGGGCCCTGGTGTATCGCAGCTG 598

QY      601  ATGGCTTGGGATATGATGATGAACCTGGTGCCTACAAACGGGCCCTGGTGTATCGCAGCTG 660
DB      601  ATGGCTTGGGATATGATGATGAACCTGGTGCCTACAAACGGGCCCTGGTGTATCGCAGCTG 660

QY      589  CTCGGATCCCAACAAGCTCTGCTGGACATGGTGGGGGGGCCCATTTGGGAGTCTTGGCG 648
DB      589  CTCGGATCCCAACAAGCTCTGCTGGACATGGTGGGGGGGCCCATTTGGGAGTCTTGGCG 648

QY      661  CTCGGATCCCAACAAGCTCTGCTGGACATGGTGGGGGGGCCCATTTGGGAGTCTTGGCG 720
DB      661  CTCGGATCCCAACAAGCTCTGCTGGACATGGTGGGGGGGCCCATTTGGGAGTCTTGGCG 720

QY      649  GGTCTCGCCTACTATTTCATGATGGGGAACCTGGGCTTAAGGTTTGATGTGACTACTC 708
DB      649  GGTCTCGCCTACTATTTCATGATGGGGAACCTGGGCTTAAGGTTTGATGTGACTACTC 708

QY      721  GGTCTCGCCTACTATTTCATGATGGGGAACCTGGGCTTAAGGTTTGATGTGACTACTC 784
DB      721  GGTCTCGCCTACTATTTCATGATGGGGAACCTGGGCTTAAGGTTTGATGTGACTACTC 784

```

Qy	709	TTTGCTCCCTAATAG	723
Db	781	TTTGTCCCTAATAG	795
RESULT 6			
ARL57325			
LOCUS	ARL57325	795 bp	DNA linear PAT 17-OCT-2001
DEFINITION	Sequence 5 from patent US 6245503.		
ACCESSION	ARL57325		
VERSION	ARL57325.1	GI:16218258	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 795)		
AUTHORS	Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.		
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use		
JOURNAL	Patent: US 6245503-A 5 12-JUN-2001;		
FEATURES	Location/Qualifiers		
source	1..795		
BASE COUNT	130 a 240 c 231 g 194 t		
ORIGIN			
Query Match	88.7%;	Score 641; DB 6; Length 795;	
Best Local Similarity	90.9%;	Pred. No. 6.3e-130;	
Matches 723; Conservative	0; Mismatches	0; Indels	72; Gaps 1;
Qy	1	ATGTTGGTAAGGTCATCGATACCCTTTACATCGCGCTTCGCCGACCTCGTGGGGGTACATT	60
Db	1	ATGTTGGTAAGGTCATCGATACCCTTTACATCGCGCTTCGCCGACCTCGTGGGGGTACATT	60
Qy	61	CCGCTCGTCGCGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Db	61	CCGCTCGTCGCGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Qy	121	CTGGAGGACGGCGTGAACTATGCACAAGGGAATTTGCCGGTTCCTCTTTCTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAACTATGCACAAGGGAATTTGCCGGTTCCTCTTTCTCTATCTTC	180
Qy	181	CTCTTGGCTTTGCTGCTGCTGTCAACCGTTCCAGCTTCAAGTATGAAGTGCACAACGTG	240
Db	181	CTCTTGGCTTTGCTGCTGCTGTCAACCGTTCCAGCTTCAAGTATGAAGTGCACAACGTG	240
Qy	241	TCCGGGTATGACATGTCACGAAACGACTGCTCCAACCTCAAGCATTTGTTATGAGGACGC	300
Db	241	TCCGGGTATGACATGTCACGAAACGACTGCTCCAACCTCAAGCATTTGTTATGAGGACGC	300
Qy	301	GACATGATCATGCACACCCC GG GTGCGTGCCTTCGCTTGGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGCACACCCC GG GTGCGTGCCTTCGCTTGGGAGAACAACTCTTCCCGC	360
Qy	361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAACGCGACGCTCCCAACACGACA	420
Db	361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAACGCGACGCTCCCAACACGACA	420
Qy	421	ATACAGCCCACTGATCGAT-----	438
Db	421	ATACAGCCCACTGATCGATTTTCTGTTGGGGCGGCTGTTTCTGTTCCGCTATGACGTG	480
Qy	439	-----TCCACAGCTGTTTCCACCATCTGCCTCGCGG	468
Db	481	GGGACCTCTCGGANCTGTCTTCTCGTCTCCACAGCTGTTTCAACATCTGCCTCGCGG	540
Qy	469	CATGAGACGGTGCAGGACTGCAAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	528
Db	541	CATGAGACGGTGCAGGACTGCAAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	600
Qy	529	ATGCTTGGGATATGATGATACTGGTCGCTTCAACGGGCCCTGGTGTATCGACGTG	588

BASE COUNT		130 a	240 c	231 g	194 t
ORIGIN					
Query Match 88.7%; Score 641; DB 6; Length 795;					
Best Local Similarity 90.9%; Pred. No. 6.3e-130;					
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;					
QY	1	ATGTTGGGTAAAGTCATCGATACCTTACATSCGGCTTCGCGACCTCGTGGGGTACATT	60		
DB	1	ATGTTGGGTAAAGTCATCGATACCTTACATSCGGCTTCGCGACCTCGTGGGGTACATT	60		
QY	61	CGCTCGTCGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATCGCTCCGGGTT	120		
DB	61	CGCTCGTCGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATCGCTCCGGGTT	120		
QY	121	CTGAGGACGGGTGAACCTATGCAACAGAGGAATTTGCCCGGTGCTTTCTCTATCTTC	180		
DB	121	CTGAGGACGGGTGAACCTATGCAACAGAGGAATTTGCCCGGTGCTTTCTCTATCTTC	180		
QY	181	CTCTTGGCTTTGCTGTCCTCTGACCGTTCCAGCTTCCGCTTATGAAATGCGCAAGTG	240		
DB	181	CTCTTGGCTTTGCTGTCCTCTGACCGTTCCAGCTTCCGCTTATGAAATGCGCAAGTG	240		
QY	241	TCCGGGATGACATGTACGAAACGACTGCTCAAACCTCAAGCAATGTGTATGAGGACGC	300		
DB	241	TCCGGGATGACATGTACGAAACGACTGCTCAAACCTCAAGCAATGTGTATGAGGACGC	300		
QY	301	GACATGATCATGACACCCCGGGTGGTCCCTGCGTTCCGGAGAAACAATCTTCCCGC	360		
DB	301	GACATGATCATGACACCCCGGGTGGTCCCTGCGTTCCGGAGAAACAATCTTCCCGC	360		
QY	361	TGCTGGGTAGCGTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCACACGACA	420		
DB	361	TGCTGGGTAGCGTACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCACACGACA	420		
QY	421	ATACGACGCCACGTCGAT-----	438		
DB	421	ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGTTCTTCTGTTCCGTTATGTACGTG	480		
QY	439	-----TCCAGCTGTTACCACTCTCGCTTCGCGCG	468		
DB	481	GGGGACCTCTCGGATCTGCTTCTCTCGTCTCCAGCTGTTACCACTCTCGCTTCGCGCG	540		
QY	469	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGT	528		
DB	541	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGT	600		
QY	529	ATGGCTTGGGATATGATGAACTGTGTGCGCTACACGGCCCTGGTATCGCAGCTG	588		
DB	601	ATGGCTTGGGATATGATGAACTGTGTGCGCTACACGGCCCTGGTATCGCAGCTG	660		
QY	589	CTCCGGATCCACAGCTGCTGTGACATGGTGGCGGGGCCCATATGGGGAGTCTCTGGCG	648		
DB	661	CTCCGGATCCACAGCTGCTGTGACATGGTGGCGGGGCCCATATGGGGAGTCTCTGGCG	720		
QY	649	GGTCTCGCTACTATTTCATGTTGGGAACTCGGGTAAAGTTTGAATGTGATGCTACTC	708		
DB	721	GGTCTCGCTACTATTTCATGTTGGGAACTCGGGTAAAGTTTGAATGTGATGCTACTC	780		
QY	709	TTTGTCTCCCTAATAG 723			
DB	781	TTTGTCTCCCTAATAG 795			
RESULT 9					
A48709 LOCUS A48709 2082 bp DNA linear PAT 07-MAR-1999					
DEFINITION Sequence 47 from Patent WO9604395.					
ACCESSION A48709					
VERSION A48709.1 GI:2302422					
KEYWORDS .					
SOURCE unidentified					
ORGANISM unidentified					

unclassified.
1 (bases 1 to 2082)
Maertens,G., Bosman,F., De M.G. and Buyse,M.
PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL
Patent: WO 9604385-A 47 15-FEB-1996;
INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
Location/Qualifiers
1..2082
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..2079
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA03146.1"
/db_xref="GI:2302423"
/translation="NLGVLDLTCGFADLVGYIPLVGAIPGGGAARALAHGVRLVLEGG
VNYATGNLPGCSFSLFLALISCLTVPASAYEVNRVSMGVHTVNDSCNSIVYEAADM
IMHTPCQVPCVNRSSNCWALPTLAARNASVPTTIRRHVDLLVGAAGFASAMTV
GLDCSVFLNSOLFTISPRRHETVQDCNSIYPGHIICHRMADWMMNWSPTTALIVS
QLLRIPQAVDVAAGHWGLAGLAYISMVGNWAKVLYVMLLFAGVUGHTRVSGAAA
SDTRGLSVLSPGSAQKIQLVNTGNSWHINLTALNCNDSLQTFGFAALFYKHFNSSG
CERLASCRSIDKFAQIGPLTYTPNSSDQPCYCHYAPRPGIVPASQGVPTVCF
TPSVNIDTRFVPTYNWGANSDVLIANNRPGRNPGCTWNNGTGFTKTCGP
PCNIGAGNNTLCTPCDFKRHPATYARCGSPWLTPRCMVHPYRLMHPYPCTVNPT
IFPKWYGVHEHREFAACNWRGCRDLEDRSELSPILLSSTENWOLILPCSPFTLP
ALSTGLIHLHVNIDVQYLVGVSAVSLVKNWEYVLLFLLLADARICACLMWMLLI
AQAEALLENLVNNAVAAGHTLSFLVFFCAAWYIKGRVLPVGAAYAFYGVWPLLLI
LLALPPRAYA"
mat_peptide
1..2076
/product="unnamed"
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN

Query Match 86.4%; Score 624.8; DB 6; Length 2082;
Best Local Similarity 90.5%; Pred. No. 2.1e-126;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;
4 TTGGGTAAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGTACATTCCG 63
4 TTGGGTAAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGTACATTCCG 63
64 CTCGTGGCGGCCCCCTAGGGGGCGCTGCGGAGGCGCTTGGCGCATGCGGCTCCGGGTTCTG 123
64 CTCGTGGCGGCCCCCTAGGGGGCGCTGCGGAGGCGCTTGGCGCATGCGGCTCCGGGTTCTG 123
124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGGCCCGGTTCCTTTCTCTATCTTCTC 183
124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGGCCCGGTTCCTTTCTCTATCTTCTC 183
184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
244 GGGATGTACATGTACGAAACGACTGTCTCAACTCAAGCATTTGTGTATGAGGACGGGAC 303
244 GGGATGTACATGTACGAAACGACTGTCTCAACTCAAGCATTTGTGTATGAGGACGGGAC 303
304 ATGATCATGTACACACCCCGGGTGGCGTTCGCTTCGGGAGAACAACTTTCCCGTGC 363
304 ATGATCATGTACACACCCCGGGTGGCGTTCGCTTCGGGAGAACAACTTTCCCGTGC 363
364 TGGGTAGCGCTACCCCCACGCTCGCAGTAGGAACGCGGCGTCCCCACACGACAAATA 423
364 TGGGTAGCGCTACCCCCACGCTCGCAGTAGGAACGCGGCGTCCCCACACGACAAATA 423
424 CGACGCCACGTCGAT----- 438
424 CGACGCCACGTCGATTTGCTTGGGGGGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
CDS

QY 439 -----TCCAGCTGTTCACCATCTCGCTTCGCGCGCAT 471
DB 484 GACCTCTGGGATCTGTCTCTCTCTCCAGCTGTTCACCATCTCGCTTCGCGCGCAT 543
QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 531
DB 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 603
QY 532 GCTTGGGATATGATGATGAACCTGGTTCGCTACAAACGGCGCTTGGTGGTATCGCAGCTGCTC 591
DB 604 GCTTGGGATATGATGATGAACCTGGTTCGCTACAAACGGCGCTTGGTGGTATCGCAGCTGCTC 663
QY 592 CGGATCCACAGCTGCTGCTGCAACATGCTGGCGGGGGCCCAATTGGGAGTCTCTGGCGGGT 651
DB 664 CGGATCCACAGCTGCTGCTGCAACATGCTGGCGGGGGCCCAATTGGGAGTCTCTGGCGGGC 723
QY 652 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAGGTTTTTGTATGATGCTACTCTTT 711
DB 724 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAGGTTTTTGTATGATGCTACTCTTT 783
QY 712 GC 713
DB 784 GC 785

RESULT 10
LOCUS AR157350 2082 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 47 from patent US 6245503.
ACCESSION AR157350
VERSION AR157350.1 GI:16218284
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2082)
AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
TITLE Purified Hepatitis C virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: US 6245503-A 47 12-JUN-2001;
FEATURES
source Location/Qualifiers
1..2082
/organism="unknown"
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN

Query Match 86.4%; Score 624.8; DB 6; Length 2082;
Best Local Similarity 90.5%; Pred. No. 2.1e-126;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;
4 TTGGGTAAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGTACATTCCG 63
4 TTGGGTAAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGTACATTCCG 63
64 CTCGTGGCGGCCCCCTAGGGGGCGCTGCGGAGGCGCTTGGCGCATGCGGCTCCGGGTTCTG 123
64 CTCGTGGCGGCCCCCTAGGGGGCGCTGCGGAGGCGCTTGGCGCATGCGGCTCCGGGTTCTG 123
124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGGCCCGGTTCCTTTCTCTATCTTCTC 183
124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGGCCCGGTTCCTTTCTCTATCTTCTC 183
184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
244 GGGATGTACATGTACGAAACGACTGTCTCAACTCAAGCATTTGTGTATGAGGACGGGAC 303
244 GGGATGTACATGTACGAAACGACTGTCTCAACTCAAGCATTTGTGTATGAGGACGGGAC 303
304 ATGATCATGTACACACCCCGGGTGGCGTTCGCTTCGGGAGAACAACTTTCCCGTGC 363
304 ATGATCATGTACACACCCCGGGTGGCGTTCGCTTCGGGAGAACAACTTTCCCGTGC 363
364 TGGGTAGCGCTACCCCCACGCTCGCAGTAGGAACGCGGCGTCCCCACACGACAAATA 423
364 TGGGTAGCGCTACCCCCACGCTCGCAGTAGGAACGCGGCGTCCCCACACGACAAATA 423
424 CGACGCCACGTCGAT----- 438
424 CGACGCCACGTCGATTTGCTTGGGGGGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483

Db 304 ATGATCATGACACACCCCGGGTGGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTCACCCCAACGCTCGACGTAAGAAAGCAGCGCTCCCAACACGACAATA 423
Db 364 TGGGTAGCGCTCACCCCAACGCTCGACGTAAGAAAGCAGCGCTCCCAACACGACAATA 423
QY 424 CGACGCCAGCTCGAT----- 438
Db 424 CGACGCCAGCTCGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGCTACGTGGG 483
QY 439 -----TCCAGCTGTTTCAACATCTCCCTCGCGGCAT 471
Db 484 GACCTCTCGGATCTGCTCTCTCCAGCTCCACAGCTGTTACCATCTCGCCTCGCGGCAT 543
QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATG 531
Db 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATG 603
QY 532 GCTTGGGATATGATGAATGAACTGCTCGCTCAACACGGCCCTGGTATGCGAGTGTCTC 591
Db 604 GCTTGGGATATGATGAATGAACTGCTCGCTCAACACGGCCCTGGTATGCGAGTGTCTC 663
QY 592 CGATCCCAACAGCTGTCGTGGAATGCTGGCGGGGCCCATTTGGGAGTCTCGCGGGGT 651
Db 664 CGATCCCAACAGCTGTCGTGGAATGCTGGCGGGGCCCATTTGGGAGTCTCGCGGGGC 723
QY 652 CTGCGCTACTATTCCATGTTGGGGAATCTGGGCTTAAGTTTGTGATGCTACTCTTT 711
Db 724 CTGCGCTACTATTCCATGTTGGGGAATCTGGGCTTAAGTTTGTGATGCTACTCTTT 783
QY 712 GC 713
Db 784 GC 785

RESULT 11
AX452796 2082 bp DNA linear PAT 06-JUL-2002
LOCUS Sequence 47 from Patent EP1211315.
DEFINITION AX452796
ACCESSION AX452796
VERSION AX452796.1 GI:21712481
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE
1 Maertens, G., Bosman, P., de Martynoff, G. and Buyse, M.A.
AUTHORS Recombinant vectors for producing hcv envelope proteins
TITLE Patent: EP 1211315-A 47 05-JUN-2002;
JOURNAL Innogenetics N.V. (BE)
LOCATION/Qualifiers

FEATURES
Source
1..2082
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
<1..2079
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38082.1"
/db_xref="GI:21712482"

CDS
1
/translation="NLGKVIDTLTCTGFADLVGYIPLVGPALGGAARALAHGVRVLEEDG
VNTATGNLPGCSFSIPLALLSCLTPASAYVRNVGMVHTWDCNSSIVYEADAM
IMHTPCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDLLVGAAPFCSAMVY
GDLICGSFVLVSQFTISPRHETVQDCNSIYPGHTGHRMADMMWNSPTTALVVS
QLIRIPQAVVAVAGHWGLAGLAVYSMVGNWAKVLVYMLLPAGVDGHTVSGGAA
SDTRGLVLSIFSPGSAQIKQLVNTGNSWHINRTALNCDSLQTFPAALFYKHKNSSG
CPRLASCRSIIKFAQGWGLTYTTPNSDQRPYCMWYAPRPGIIVPASQVCGPYICF
TPSPVVTGTRDGVPTYNWGNDSVLILNNRPPRGNWFCQTNWGTGKTCGGP
PCNIGAGNNTLCTDPCRKHPEATYARCQSGPWLTPRCWVHYPIRLWHYPTVNTF
IFKRMVGVGVEHFEACNWTGECRDLDRSELSPLLLSTTEWQLIPCSFTTLP
ALSTGLIHLHQNIIVDQVLYGVGSVAVSLVIKWEYVLLFLALLADARACILMMMLLI
AQAEALLENLVILNAAAVAGAGHTLSFLVFFCNAMVIKGRILVPGAAAYFYGVWPLLL

mat_peptide 1..2076 /product="unnamed"
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN
Query Match 86.4%; Score 624.8; DB 6; Length 2082;
Best Local Similarity 90.5%; Pred. No. 2.1e-126;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;
QY 4 TTGGGTAAAGTTCATCGCATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCG 63
Db 4 TTGGGTAAAGTTCATCGCATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCG 63
QY 64 CTGCTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCG 123
Db 64 CTGCTCGGCGCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCG 123
QY 124 GAGGACGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTTTCTCTATCTCTC 183
Db 124 GAGGACGGCTGAACATATGCAACAGGGAATTTGCCCGGTTGCTTTCTCTATCTCTC 183
QY 184 TTGGCTTTGCTGCTGTCTGACCGTTCCAGCTTTCGCTTATGAAGTGGCAACGTGTCC 243
Db 184 TTGGCTTTGCTGCTGTCTGACCGTTCCAGCTTTCGCTTATGAAGTGGCAACGTGTCC 243
QY 244 GGGATGATGACACACCCCGGGTGGTCCCTCGCTCAAGCAATTTGTATGAGGAGCGGAC 303
Db 244 GGGATGATGACACACCCCGGGTGGTCCCTCGCTCAAGCAATTTGTATGAGGAGCGGAC 303
QY 304 ATGATCATGACACACCCCGGGTGGTCCCTCGCTCAAGCAATTTGTATGAGGAGCGGAC 363
Db 304 ATGATCATGACACACCCCGGGTGGTCCCTCGCTCAAGCAATTTGTATGAGGAGCGGAC 363
QY 364 TGGGTAGCGCTCACCCCAACGCTCGACGTAAGAAAGCAGCGCTCCCAACGACAATA 423
Db 364 TGGGTAGCGCTCACCCCAACGCTCGACGTAAGAAAGCAGCGCTCCCAACGACAATA 423
QY 424 CGACGCCAGCTCGAT----- 438
Db 424 CGACGCCAGCTCGATTTGCTGTTGGGGGGGCTGCTTTCTGTTCCGCTATGCTAGTGGG 483
QY 439 -----TCCAGCTGTTTCAACATCTCGCCTCGCGGCAT 471
Db 484 GACCTCTCGGATCTGCTTCTCTCCAGCTTTCACCATCTCGCCTCGCGGCAT 543
QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATG 531
Db 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATG 603
QY 532 GCTTGGGATATGATGAATGAACTGCTCGCTCAACACGGCCCTGGTATGCGAGTGTCTC 591
Db 604 GCTTGGGATATGATGAATGAACTGCTCGCTCAACACGGCCCTGGTATGCGAGTGTCTC 663
QY 592 CGGATCCCAACAGCTGTCGTGGAATGCTGGCGGGGCCCATTTGGGAGTCTCGCGGGGT 651
Db 664 CGGATCCCAACAGCTGTCGTGGAATGCTGGCGGGGCCCATTTGGGAGTCTCGCGGGGC 723
QY 652 CTGCGCTACTATTCCATGTTGGGGAATCTGGGCTTAAGTTTGTGATGCTACTCTTT 711
Db 724 CTGCGCTACTATTCCATGTTGGGGAATCTGGGCTTAAGTTTGTGATGCTACTCTTT 783
QY 712 GC 713
Db 784 GC 785

RESULT 12
AX685048 2082 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 47 from Patent WO0205548.
DEFINITION AX685048
ACCESSION AX685048
VERSION AX685048.1 GI:29371453

KEYWORDS	Hepatitis C virus
SOURCE	Hepatitis C virus
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE	1
AUTHORS	Maertens, G., Bosman, F., and Buysse, M.A.
TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL	Patent: WO 02055548-A 47 18-JUL-2002;
FEATURES	INNOGENETICS N.V. (BE)
source	Location/Qualifiers
	1..2082
	/organism="Hepatitis C virus"
	/mol_type="genomic DNA"
	/db_xref="taxon:11103"
CDS	1..2079
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD86535.1"
	/db_xref="GI:29371454"
	/translation="NLGKVIDLTLCGFADLVGYIPLVGAPLGGAAALAHGVRVLEDD VNYATCNLPFGCSFIFLLALLSCLTPASAYEVNVSVMYHVTNDCSNSIYVEAADM IMHTPCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDLLVGAAPCSAMVY GDLCGSFLVLSOLFILSPRHETVQDCNCSIYPGHIHTRHMAWMMNWSPTALVVS QLLRIPQAVVDMVAGHWGLAGLAYISVMGNWAKVLVVMLLFAGVDGHRVSGGAAA SDRGLVSLFSPGSAOKIQLVNTNGSWHINRTALNCNDSLOTGFFAALFYRKHFNSSG CPERLASCRSIDKFAQGWGPLYTTPNSDQRPYCHVAPRPGIYPASQVCGPVYCF TSPVVGVTIDRFQVNTNGANDSDVILINRTPRGNWFGCTWMNGFTKTCGGP PCNIGAGNNTLTCTDCFRKHEATYARCQSPWLTPTPCMVHYVYRLWHVYCTVNT IFKVMYVGGVHEHFEACNWTGRCDLDRSELSPLLSSTWQLPCSPFTLP ALUSTGLIHLHONIVDVOYLVGVSAYVSLVIRKVEYVLLFLLLADARICACUWMLLI AQEAALLENVLVINAAGVAGAGHTLSFLVFFCAAWYIKRLVPGAAYAYFYGVWPLLL LLLALPPRAYA"
mat_peptide	1..2076
	/product="unnamed"
BASE COUNT	366 a 634 c 600 g 482 t
ORIGIN	
Query Match	86.4%; Score 624.8; DB 6; Length 2082;
Best Local Similarity	90.5%; Pred. No. 2.1e-126;
Matches	708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;
QY	4 TTGGGTAGGTCATCATGATACCTTACATCGCGCTTCGCGCACTCGTGGGTACATTCCG 63
Db	4 TTGGGTAGGTCATCATGATACCTTACATCGCGCTTCGCGCACTCGTGGGTACATTCCG 63
QY	64 CTCGTGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGATGGCGTCCGGTTCTG 123
Db	64 CTCGTGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGATGGCGTCCGGTTCTG 123
QY	124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTCCTC 183
Db	124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTCCTC 183
QY	184 TTGGCTTTGCTGCTGCTGTCACCGTTCCAGCTTCGCTTATGAAGTGCGCAACGTGCC 243
Db	184 TTGGCTTTGCTGCTGCTGTCACCGTTCCAGCTTCGCTTATGAAGTGCGCAACGTGCC 243
QY	244 GGGATGTACCATGTACAGCAACACTGCTCAACTCAAGCATTTGTATGAGGACGGGAC 303
Db	244 GGGATGTACCATGTACAGCAACACTGCTCAACTCAAGCATTTGTATGAGGACGGGAC 303
QY	304 ATGATCATGCACACCCCGGGTGGCGCTTCGTTTCGGGAGAACAACTCTTCCCGCTGC 363
Db	304 ATGATCATGCACACCCCGGGTGGCGCTTCGTTTCGGGAGAACAACTCTTCCCGCTGC 363
QY	364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAAGCCAGCGTCCCGACACGACAATA 423
Db	364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAAGCCAGCGTCCCGACACGACAATA 423
QY	424 CGACGCCAGCTCGAT

mat_peptide
1..2427
/product="unnamed"

BASE COUNT
434 a 745 c 714 g 540 t

Db	424 CGACGCCAGCTCGATTTGCTCGTGGGGCGGCTGCTTTCTGTTCCTGATGACGTGGGG 483
QY	439 -----TCCAGCTGTTTACCACATCTCGCCTCGCCGGCAT 471
Db	484 GACCTCTCGGATCTGCTCTCTCGTCTCCAGCTGTTTACCACATCTCGCCTCGCCGGCAT 543
QY	472 GAGACGGTCGACGACTGCAATTTGCTCAATCTATCCGGCCACATACGGGTACCGTATG 531
Db	544 GAGACGGTCGACGACTGCAATTTGCTCAATCTATCCGGCCACATACGGGTACCGTATG 603
QY	532 GCTTGGGATATGATGATGAACCTGCTCGCTCAACAGGGCCCTGCTGATTCGAGCTGCTC 591
Db	604 GCTTGGGATATGATGATGAACCTGCTCGCTCAACAGGGCCCTGCTGATTCGAGCTGCTC 663
QY	592 CGGATCCCAAGCTGCTCGTGACATGTGTGGGGGGCCCATTTGGGGAGTCTCTGGCGGGT 651
Db	664 CGGATCCCAAGCTGCTCGTGACATGTGTGGGGGGCCCATTTGGGGAGTCTCTGGCGGGC 723
QY	652 CTCGCTACTATTTCATGTTGGGGAACCTGGGCTAAGGTTTTCATTTGATGATCTACTCTTT 711
Db	724 CTCGCTACTATTTCATGTTGGGGAACCTGGGCTAAGGTTTTCATTTGATGATCTACTCTTT 783
QY	712 GC 713
Db	784 GC 785
RESULT	13
LOCUS	A48711
DEFINITION	Sequence 49 from Patent WO9604385.
ACCESSION	A48711
VERSION	A48711.1
KEYWORDS	GI:2302424
SOURCE	unidentified
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 2433)
AUTHORS	Maertens, G., Bosman, F., De, M.G. and Buysse, M.
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL	Patent: WO 9604385-A 49 15-FEB-1996;
COMMENT	INNOGENETICS NV (BE)
FEATURES	Other publication CA 2172273 960215
	Other publication AU 3382495 960304.
source	Location/Qualifiers
	1..2433
	/organism="unidentified"
	/mol_type="genomic DNA"
	/db_xref="taxon:32644"
CDS	1..2430
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAA03147.1"
	/db_xref="GI:2302425"
	/translation="NSTNPKPQRTKTRNTRPQDKFPGGGQIVGGVYLLPRRGPRL GVARPKTSERSQPRGRQPIPKARRPEGRAWQPCYPMPLYNEMGMWAGWLLSPRG SPSPGPTDPRRSRLKQIDITLTCGFADLVGYIPLVGAPLGGAAALAHGVRVLEDD GNYATCNLPFGCSFIFLLALLSCLTPASAYEVNVSVMYHVTNDCSNSIYVEAAD VMHTPGVPCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDLLVGAAPCSAMV YDLCGSFLVLSOLFILSPRHETVQDCNCSIYPGHIHTRHMAWMMNWSPTALVVS SDRGLVSLFSPGSAOKIQLVNTNGSWHINRTALNCNDSLOTGFFAALFYRKHFNSS ASDTGLVSLFSPGSAOKIQLVNTNGSWHINRTALNCNDSLOTGFFAALFYRKHFNSS GCPERLASCRSIDKFAQGWGPLYTTPNSDQRPYCHVAPRPGIYPASQVCGPVYCF TSPVVGVTIDRFQVNTNGANDSDVILINRTPRGNWFGCTWMNGFTKTCGGP PCNIGAGNNTLTCTDCFRKHEATYARCQSPWLTPTPCMVHYVYRLWHVYCTVNT IFKVMYVGGVHEHFEACNWTGRCDLDRSELSPLLSSTWQLPCSPFTLP ALUSTGLIHLHONIVDVOYLVGVSAYVSLVIRKVEYVLLFLLLADARICACUWMLLI AQEAALLENVLVINAAGVAGAGHTLSFLVFFCAAWYIKRLVPGAAYAYFYGVWPLLL LLLALPPRAYA"
mat_peptide	1..2427
	/product="unnamed"
BASE COUNT	434 a 745 c 714 g 540 t

REFERENCE 1 (bases 1 to 2433)									
AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.									
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use									
JOURNAL Patent: US 6245503-A 49 12-JUN-2001;									
FEATURES Location/Qualifiers									
source 1..2433									
BASE COUNT 434 a 745 c 714 g 540 t									
ORIGIN /organism="unknown"									
Query Match 86.4%; Score 624.8; DB 6; Length 2433;									
Best Local Similarity 90.5%; Pred. No. 2.1e-126;									
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;									
QY	4	TTGGGTAAGGTCAATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATTCCG	63						
Db	355	TTGGGTAAGGTCAATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATTCCG	414						
QY	64	CTCGTCGGCGCCCTAGGGGCGCTGCCAGGCGCTTCGGGCATGGCGTTCGGTTCGT	123						
Db	415	CTCGTCGGCGCCCTAGGGGCGCTGCCAGGCGCTTCGGGCATGGCGTTCGGTTCGT	474						
QY	124	GAGGACGGGTGAACCTATATGCAACAGGGAATTGCCCGGTTCCTTCTATCTTCCTC	183						
Db	475	GAGGACGGGTGAACCTATATGCAACAGGGAATTGCCCGGTTCCTTCTATCTTCCTC	534						
QY	184	TTGGCTTTGCTGTCCTCTGACCGTTCCAGCTTCGCGCTTATGAAGTGCACACGTGTC	243						
Db	535	TTGGCTTTGCTGTCCTCTGACCGTTTCAGCTTCGCGCTTATGAAGTGCACACGTGTC	594						
QY	244	GGGATGTACCATGTACAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGGACGGAC	303						
Db	595	GGGATGTACCATGTACAGAACGACTGCTCAACTCAAGCATTTGTGTATGAGGACGGAC	654						
QY	304	ATGATCATGACACCCCGGGTGGCTGCGCTCGGAGAACAACTCTTCCCGCTGC	363						
Db	655	ATGATCATGACACCCCGGGTGGCTGCGCTCGGAGAACAACTCTTCCCGCTGC	714						
QY	364	TGGGTAGCGCTCACCCGCCACGCTCGAGCTAGGAAACGCGCTCCACACGACAATA	423						
Db	715	TGGGTAGCGCTCACCCGCCACGCTCGAGCTAGGAAACGCGCTCCACACGACAATA	774						
QY	424	CGACGCCACGTCGAT	438						
Db	775	CGACGCCACGTCGATTTTGTCTGTTGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGG	834						
QY	439	-----TCCAGCTGTTTCAACATCTCGCTCGCGCGGCAAT	471						
Db	835	GACCTCTCGGATCTGTCTTCTCTCCGCTCCAGCTGTTTCAACATCTCGCTCGCGCGCAAT	894						
QY	472	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATG	531						
Db	895	GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTACCGGTATG	954						
QY	532	GCTTGGGATATGATGATGAACCTGCTGCGCTACAAACGGCCCTGGTGTATCGCAGTGTCT	591						
Db	955	GCTTGGGATATGATGATGAACCTGCTGCGCTACAAACGGCCCTGGTGTATCGCAGTGTCT	1014						
QY	592	CGGATCCCAACAGCTGCTGGACATGTTGGCGGGGCCCATTTGGGAGTCTTGGGGGT	651						
Db	1015	CGGATCCCAACAGCTGCTGGACATGTTGGCGGGGCCCATTTGGGAGTCTTGGGGGT	1074						
QY	652	CTCGCTACTATTCCATGTTGGGAACTGGGCTTAAGTTTGTATGATGCTACTCTTT	711						
Db	1075	CTCGCTACTATTCCATGTTGGGAACTGGGCTTAAGTTTGTATGATGCTACTCTTT	1134						
QY	712	GC 713							
Db	1135	GC 1136							
RESULT 14									
AR157351									
LOCUS 2433 bp DNA linear PAT 17-OCT-2001									
DEFINITION Sequence 49 from patent US 6245503.									
ACCESSION AR157351									
VERSION AR157351.1 GI:16218285									
KEYWORDS Unknown.									
SOURCE Unknown.									
ORGANISM Unclassified.									

AX452798	AX452798	2433 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	Sequence 49 from Patent EPI211315.				
DEFINITION	AX452798				
ACCESSION	AX452798.1	GI:21712483			
VERSION					
KEYWORDS	Hepatitis C virus				
SOURCE	Hepatitis C virus				
ORGANISM	Hepatitis C virus				
REFERENCE	1				
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.				
TITLE	Recombinant vectors for producing hcv envelope proteins				
JOURNAL	Patent: EP 1211315-A 49 05-JUN-2002;				
FEATURES	Innogenetics N.V. (BE)				
source	Location/Qualifiers				
	1..2433				
	/organism="Hepatitis C virus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:11103"				
CDS	1..2430				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD38083.1"				
	/db_xref="GI:21712484"				
	/translation="MSTNPKPQKTKNTRRNPQDVKPPGGQIVGGVLLPRRPRRL GVARTKTSEROPRRROPIPKARPEGRAWAQPGYPWPLYGNEGMGWAGLLSPRG SRSPGPTDPRRSRLNGLVIDLTGCFADLVGYIPLVAGPLGGAARALAHGVRVLED GVNYATNLPKGSFSLFLALLSLCTVPASAVEVRNVGMVHVTNDCSNSSIVYBAAD MIMHTPGVCVPCRENNSSRCWALPTLAARNASVPTTIRHVDLLVGAAPFCSAMV VGDLCGSPLVLSQLFTISPRRHETVQDCNSIYPGHITGRMANWDMNWSPTTALIV SQLLRIPQAVDMVAGHVLGVLAVSMVGNWAKVLVLMVLLFAGVDGHTRVSGAA ASDTRGLVSLFSPGSAQKTQLVNTNGSWHINRTALNCNDSLQTFPAALFYKHFNSS GCEPLASCRSDKFAQMGPLTYTEPNSSDQRCWYAPRPGVIGVPAQVCGPVYC FTPSVVTGTRDFGVPTVNWGANDSDVILNNTPRGNWFGCTWMNGTGKTCGG PFCNIGAGNNLTCTPTDCFRKHPEATYACGSGPWLTPRCMVHYPRLYRHYVCTVNF TIPKVMYVGGVEHPEAACNTRGERCDLEDRDRSELSPLLLSTTEWQILPSCFTTL PALSTGLIHLHQNIVDVQYLVGSAVSLVIRKWEYVLLFLLLADARICACIMMMLL IAQAEAALENLVLVNAAAVAGAHGTLFLVFFCAAWYIKGRILVPGAAYAFYGVWPLL LLALLPPRAYA"				
	1..2427				
mat_peptide	/product="unnamed"				
BASE COUNT	434 a 745 c 714 g 540 t				
ORIGIN					
	Query Match 86.4%; Score 624.8; DB 6; Length 2433;				
	Best Local Similarity 90.5%; Pred. No. 2.1e-126;				
	Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;				
QY	4 TTGGGTAAGGTTCATCGATACCTTACATCGGCTTCGCGACCTCGTGGGGTACATTCCG 63				
DB	355 TTGGGTAAGGTTCATCGATACCTTACATCGGCTTCGCGACCTCGTGGGGTACATTCCG 414				
QY	64 CTGCTCGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGGACCTCGTGGGGTACATTCCG 123				
DB	415 CTGCTCGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGGACCTCGTGGGGTACATTCCG 474				
QY	124 GAGGACGGGTGAACATATGCAACAGGAAATTTGCCCGGTTGCTCTCTATCTTCCTC 183				
DB	475 GAGGACGGGTGAACATATGCAACAGGAAATTTGCCCGGTTGCTCTCTATCTTCCTC 534				
QY	184 TTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCGCGCTTATGAAGTGGCAACGTGTCC 243				
DB	535 TTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCGCGCTTATGAAGTGGCAACGTGTCC 594				
QY	244 GGGATGTACATGTCAAGAAAGATGCTCAACTCAAGCATTTGTATGAGGAGCGGAC 303				
DB	595 GGGATGTACATGTCAAGAAAGATGCTCAACTCAAGCATTTGTATGAGGAGCGGAC 654				
QY	304 ATGATCATGCACACCCCGGTCGCTGCGTTCGGGAGAACACTCTTCCCGCTGC 363				
DB	655 ATGATCATGCACACCCCGGTCGCTGCGTTCGGGAGAACACTCTTCCCGCTGC 714				

Search completed: December 20, 2003, 02:02:00
Job time : 2742.48 secs

QY	364 TGGGTAGCGCTCACCCCAAGCTCGCAGCTAGCAAGCGCGCTCCCAACGACAAATA 423				
DB	715 TGGGTAGCGCTCACCCCAAGCTCGCAGCTAGCAAGCGCGCTCCCAACGACAAATA 774				
QY	424 CGACGCCACGTCGAT-----TCCCAGCTGTTTCCCGCTATGATGCTGCGGG 438				
DB	775 CGACGCCACGTCGATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGATGCTGCGGG 834				
QY	439 -----TCCCAGCTGTTTCCCGCTATGATGCTGCGGGCAT 471				
DB	835 GACCTCTGCGGATCTGTTCTTCCGCTCCAGCTGTTTCCCGCTATGATGCTGCGGGCAT 894				
QY	472 GAGACGGTCCAGGACTGCAATTGCTCAATCTATCCCGGCAATACCGGCTACCGGTATG 531				
DB	895 GAGACGGTCCAGGACTGCAATTGCTCAATCTATCCCGGCAATACCGGCTACCGGTATG 954				
QY	532 GCTTGGGATATGATGATGAACTGGTCCCTACAACGGCGCTGGTGGTATCGCAGCTGCTC 591				
DB	955 GCTTGGGATATGATGATGAACTGGTCCCTACAACGGCGCTGGTGGTATCGCAGCTGCTC 1014				
QY	592 CGGATCCCAAGCTGTTCTGGAACATGTTGGCGGGGCGCCATTGGGGAGTCTCTGCGGGT 651				
DB	1015 CGGATCCCAAGCTGTTCTGGAACATGTTGGCGGGGCGCCATTGGGGAGTCTCTGCGGGT 1074				
QY	652 CTGCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 711				
DB	1075 CTGCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 1134				
QY	712 GC 713				
DB	1135 GC 1136				

THIS PAGE BLANK (USPTO)

Title: GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: December 19, 2003, 16:53:58 ; Search time 199.88 Seconds
(without alignments)
9764.351 Million cell updates/sec

US-09-899-303a-21
Perfect score: 723
Sequence: 1 ATGTTGGTAAAGTCACTCA.....TACTCTTGTCTCCCTAATAG 723

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_19Jun03.*			
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*				2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*			
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*				4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*			
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*				6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*			
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*				8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*			
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*				10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*			
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*				12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*			
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*				14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*			
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*				16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*			
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*				18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*			
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*				20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*			
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*				22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*			
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*				24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*			
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*							
SUMMARIES							
Result No.	Score	Query Match	Length	DB	ID	Description	
1	723	100.0	723	17	AAT12961	HCV E1 construct H	
2	723	100.0	723	24	AAL48925	Hepatitis C virus	
3	641	88.7	795	17	AAT12705	HCV E1 construct H	
4	641	88.7	795	24	AAL48914	Hepatitis C virus	
5	624.8	86.4	2082	24	AAL48939	Hepatitis C virus	
6	624.8	86.4	2086	17	AAT12973	HCV E1 construct H	
7	624.8	86.4	2433	17	AAT12974	HCV E1 construct H	
8	613.8	84.9	2434	24	AAL48940	Hepatitis C virus	

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

9	598.2	82.7	606	17	AAT12963	HCV E1 construct H
10	598.2	82.7	606	24	AAL48927	Hepatitis C virus
11	597	82.6	636	17	AAT12964	HCV E1 construct H
12	597	82.6	636	24	AAL48928	Hepatitis C virus
13	559.2	77.3	9605	24	ABK91411	Hepatitis C virus
14	559.2	77.3	9605	24	ABK91424	Hepatitis C virus
15	559.2	77.3	9605	24	ABK91425	Hepatitis C virus
16	559.2	77.3	9605	24	ABK91426	Hepatitis C virus
17	559.2	77.3	9605	24	ABK91428	Hepatitis C virus
18	559.2	77.3	9605	24	ABK91429	Hepatitis C virus
19	559.2	77.3	9605	24	ABK91430	Hepatitis C virus
20	559.2	77.3	9605	24	ABK91431	Hepatitis C virus
21	559.2	77.3	9605	24	ABK91432	Hepatitis C virus
22	559.2	77.3	9605	24	ABK91433	Hepatitis C virus
23	559.2	77.3	9605	24	AAD25332	Hepatitis C virus
24	559.2	77.3	9608	24	ABK91427	Hepatitis C virus
25	559.2	77.3	11062	24	AAD25331	Hepatitis C virus
26	559.2	77.3	11076	21	AAA98965	Hepatitis C virus
27	556	76.9	561	17	AAT12962	HCV E1 construct H
28	556	76.9	561	24	AAL48926	Hepatitis C virus
29	553.8	76.6	1251	13	AAQ26981	HCV gene 1. Hepat
30	553.8	76.6	3360	17	AAT03677	Hepatitis C genome
31	553.8	76.6	3461	15	AAQ64068	Non-A, non-B hepat
32	553.8	76.6	3461	16	AAT30386	S'UTR/CORE/ENV/NS1
33	553.8	76.6	9413	16	AAT03960	Partial HCV non-st
34	553.8	76.6	9413	16	AAQ81559	Hepatitis C virus
35	553.8	76.6	9413	16	AAQ80498	DNA encoding HCV p
36	553.8	76.6	9413	24	AAD25517	Hepatitis C virus
37	553.8	76.6	9413	25	AAD49655	Hepatitis C virus
38	553.8	76.6	9413	25	AAL53723	Hepatitis C virus
39	552.8	76.5	1562	19	AAV60672	Fragment #5 isolat
40	552.8	76.5	1953	25	AAAL5222	Plasmid pIDK2 DNA
41	552.8	76.5	2187	19	ABA03491	Cuticic protein 1
42	552.8	76.5	2540	14	AAQ43889	NANB hepatitis vir
43	552.8	76.5	2540	15	AAQ63753	NANBH genomic fra
44	552.8	76.5	2829	19	AAV60673	Fragment #6 isolat
45	552.2	76.4	9609	24	AAD33038	HCV-S1 full-length

ALIGNMENTS

RESULT 1	
AAT12961	
ID	AAT12961 standard; DNA; 723 BP.
XX	AC AAT12961;
XX	AC AAT12961;
DT	24-SEP-1996 (first entry)
DE	HCV E1 construct HCCI37.
XX	HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX	ss.
OS	Hepatitis C virus.
XX	XX
PN	WO9604385-A2.
XX	15-FEB-1996.
PD	31-JUL-1995; 95WO-EP03031.
XX	29-JUL-1994; 94EP-0870132.
XX	(INNO-) INNOGENETICS NV.
PA	Bosman F, Buyse M, De Martynoff G, Maertens G;
XX	WPI; 1996-129401/13.
DR	Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
XX	PT

proteins - in presence of di: sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

Claim 23: Fig 21: 146pp; English.

AA12704-T12709 and AA12361-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for *in vitro* detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

BP: 126 A: 220 C: 208 G: 169 T: 0 other;

Query Match	100.0%	Score 723	DB 17	Length 723
Best Local Similarity	100.0%	Pred. No. 2.2e-185		
Conservative	0	Mismatches	0	Gaps 0

QY	1	ATGTTGGGTAAGGT	CATCGATACCC	TTACATGCGG	TTCCGCGA	CCTCGTGGGT	TACATT	60
Db	1	ATGTTGGGTAAGGT	CATCGATACCC	TTACATGCGG	TTCCGCGA	CCTCGTGGGT	TACATT	60
QY	61	CGCGTCGT	CGGGCCGCC	CTTAGGGGGCG	CTGCAGGGCC	CTGCGCAT	GGGTCCGGGT	120
Db	61	CGCGTCGT	CGGGCCGCC	CTTAGGGGGCG	CTGCAGGGCC	CTGCGCAT	GGGTCCGGGT	120
QY	121	CTTGAGGAC	GGCGTGA	CTATGCAAC	AGGAAATT	GC	CCGGTGTCTTTCTCTATCTTC	180
Db	121	CTTGAGGAC	GGCGTGA	CTATGCAAC	AGGAAATT	GC	CCGGTGTCTTTCTCTATCTTC	180
QY	181	CTCTTGGCT	TGTGTCCT	GTCAGCG	TTCCAGCT	TCCGCT	TATGAACTGGCAACGTG	240
Db	181	CTCTTGGCT	TGTGTCCT	GTCAGCG	TTCCAGCT	TCCGCT	TATGAACTGGCAACGTG	240
QY	241	TCGGGATGT	TACCATGT	CAGAA	CGACTGT	CCAACT	CAAGCATTTGTATGAGCGAGCG	300
Db	241	TCGGGATGT	TACCATGT	CAGAA	CGACTGT	CCAACT	CAAGCATTTGTATGAGCGAGCG	300
QY	301	GACATGAT	CATCCA	CACCCCGGG	TGCGTGC	CTTCGGGAG	AACAATCTTCCCGC	360
Db	301	GACATGAT	CATCCA	CACCCCGGG	TGCGTGC	CTTCGGGAG	AACAATCTTCCCGC	360
QY	361	TGCTGGGT	AGCGCT	CACCCCA	CGCTCCGAC	TAGGAAC	CGCAGCGTCCCCACCA	420
Db	361	TGCTGGGT	AGCGCT	CACCCCA	CGCTCCGAC	TAGGAAC	CGCAGCGTCCCCACCA	420
QY	421	ATACGAC	GGCAC	CGTGGATT	CCAGCT	GTTCAC	ATCTCGCTTCGCGGCATGAGCGGTG	480
Db	421	ATACGAC	GGCAC	CGTGGATT	CCAGCT	GTTCAC	ATCTCGCTTCGCGGCATGAGCGGTG	480
QY	481	CAGGACT	GTGAA	TGTCTCA	ATCTAT	CCCGGCC	CACATAACCGGTTCACCGTATGGCTTGGGAT	540
Db	481	CAGGACT	GTGAA	TGTCTCA	ATCTAT	CCCGGCC	CACATAACCGGTTCACCGTATGGCTTGGGAT	540
QY	541	ATGATGAT	GAACTGG	TCGGCT	CAAGCG	CCCTGGTGG	TATCGGAGCTCTCTCCGGATCCCA	600
Db	541	ATGATGAT	GAACTGG	TCGGCT	CAAGCG	CCCTGGTGG	TATCGGAGCTCTCTCCGGATCCCA	600
QY	601	CAAGCT	TCGTGGAC	ATGGTGG	CGGGGG	CCCATTC	GGGAGTCTTGGGGGGTCTCGCCTAC	660

601	CAAGCTGTCGTGCACATGCTGGCGGGGGCCCATTCGGGAGTCTCTGGCGGCTCTCGCGCTAC	720
661	TATTCACATGCTGGGAACTCGGCTTAAGGTTTTCATTGTGATGCTACTCTTTGCTCCCTAA	720

721 TAG 723

DB 721 TAG 723

RESULT 2

AA148925
ID AA1.48925 standard: DNA: 723 BP.

XX
30 MAY 1967E.XX
24-OCT-20[illegible]XX
F1 protein: E2 protein: infection;

virucide; immunostimulant; vaccine; ds.

OS Hepatitis C virus.

AA
PN
WO20025548-A2.

XX
PD
18-JUL-2002.XX
11-TAN-2002:XX
11 TAN-2001: 2001IS-260699P.

PR 30-AUG-2001; 2001US-315768P.

PA (INNO-) INNOGENETICS NV.

PI Maertens G, Bosman F, Buyse M;

DR WPI; 2002-599657/64.

DR F-PSDB/ AAC10000;
XX

PT New therapeutic vaccine compositions
PT recombinant hepatitis C virus (HCV) single or specific oligomeric

PT recombinant envelope
from HCV infection

XX 173-174. 243pp. English.

XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC

XX sequence 723 BP; 126 A; 220 C; 208 G; 169 T; 0 other;
 SO

Query Match	100.0%;	Score 723;	DB 24;	Length 723;
-------------	---------	------------	--------	-------------

Best Local Similarity	100.0%	Freq. No. Errors	0;
Mismatches	0;	Mismatches	0;
Indels	0;	Indels	0;
Gaps	0;	Gaps	0;

1 ATGTTGGTAAGTTCATCGGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

1 TCTTCCCTACCTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

C: CCGGTCCGACGGGGCGCTGCCAGGGCCCCTGGCGCATGGCGTCCGGTT 120


```
Db 61 CCGCTGTCGGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCGGGGT 120
Qy 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTTCTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTTCTTCTATCTTC 180
Qy 181 CTCCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTCCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy 361 TCGTGGGTAGCGCTCACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 420
Db 361 TCGTGGGTAGCGCTCACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 420
Qy 421 ATACGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db 421 ATACGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACCGGCTGATGATGATGATGATGATG 540
Db 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACCGGCTGATGATGATGATGATGATG 540
Qy 541 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 601 CAAGCTGCTGGACATGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Db 601 CAAGCTGCTGGACATGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Qy 661 TATTCATGCTGGGGAACCTGGGCTAAGGTTTGAATGATGATGATGATGATGATGATGATG 720
Db 661 TATTCATGCTGGGGAACCTGGGCTAAGGTTTGAATGATGATGATGATGATGATGATGATG 720
Qy 721 TAG 723
Db 721 TAG 723
RESULT 3
AAT12705
ID AAT12705 standard; DNA; 795 BP.
XX AC
AC AAT12705;
XX 23-SEP-1996 (first entry)
XX HCV E1 construct HCC110A.
DE DE
DE HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
OS Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-BP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
PA
```

```
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;
```

```
Query Match 88.7%; Score 641; DB 17; Length 795;
Best Local Similarity 90.9%; Pred. No. 3.2e-163;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
Qy 1 ATGTTGGGTAAGTCAATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CCGCTGTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 CCGCTGTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTTCTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTTCTTCTATCTTC 180
Qy 181 CTCCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTCCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy 361 TCGTGGGTAGCGCTCACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 420
Db 361 TCGTGGGTAGCGCTCACCCCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 420
Qy 421 ATACGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 438
Db 421 ATACGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy 439 -----TCCGAGCTGTTTACCACCTCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 468
Db 481 GGGGACCTCTCGGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
```

469 CATGAGCGTCCAGGCTCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGT 528
541 CATGAGCGTCCAGGCTCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGT 600
529 ATGGCTTGGGATATGATGATGAATCGTTCGCTACAAACGGCCCTGGTGGTATCCAGCTG 588
601 ATGGCTTGGGATATGATGATGAATCGTTCGCTACAAACGGCCCTGGTGGTATCCAGCTG 660
589 CTCGGATCCCAAGCTGCTGTGACATGTTGGGCGGGGCCCATTTGGGAGTCTCTGGCG 648
661 CTCGGATCCCAAGCTGCTGTGACATGTTGGGCGGGGCCCATTTGGGAGTCTCTGGCG 720
649 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTC 708
721 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTC 780
709 TTTGCTCCCTAATAG 723
781 TTTGCTCCCTAATAG 795
RESULT 4
ID AAL48914
XX AAL48914 standard; DNA; 795 BP.
AC AAL48914;
XX
DT 24-OCT-2002 (first entry)
XX Hepatitis C virus clone HCC110A E1 protein coding sequence.
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX
XX
XX
XX WO200255548-A2.
XX
XX 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
PR 30-AUG-2001; 2001US-315768P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buyse M;
XX
XX WPI; 2002-599657/64.
DR P-PSDB; AAO18661.
XX
XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX
XX Example 2; Page 161-162; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosis. The present sequence
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 88.7%; Score 641; DB 24; Length 795;
Best Local Similarity 90.9%; Pred. No. 3.2e-163;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAAGTCAATGATACCTTATATGCGGCTTCGCGAGACTGTGGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCAATGATACCTTATATGCGGCTTCGCGAGACTGTGGGGTACATT 60
QY 61 CGCTCGTTCGGGCGCCCTTAGGGGCGCTGCAGGGCCCTGCGCATGCGGCTTCGGGTT 120
Db 61 CGCTCGTTCGGGCGCCCTTAGGGGCGCTGCAGGGCCCTGCGCATGCGGCTTCGGGTT 120
QY 121 CTGAGGAGCGGTGAATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTC 180
Db 121 CTGAGGAGCGGTGAATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGTGACCGTTCCAGCTTCGCTTATGAGTGGCGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTGTGACCGTTCCAGCTTCGCTTATGAGTGGCGCAACGTG 240
QY 241 TCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAG 300
Db 241 TCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAG 300
QY 301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTTTCCCGC 360
Db 301 GACATGATCATGACACCCCGGGTGGTCCCTGGTTCGGGAGAACAACTTTCCCGC 360
QY 361 TGTGCTGAGGCTACCCCGGCTGCGAGTGGAGAACCGGCTCCCAACGACGACA 420
Db 361 TGTGCTGAGGCTACCCCGGCTGCGAGTGGAGAACCGGCTCCCAACGACGACA 420
QY 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG 480
Db 421 ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG 480
QY 439 -----TCCAGCTGTTCCACCATCTCGCTCGCGG 468
Db 481 GGGGACCTCTGGGATCTGTCTTCTCTGCTCCAGCTGTTCCACCATCTCGCTCGCGG 540
QY 469 CATGAGAGGTGCGAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGT 528
Db 541 CATGAGAGGTGCGAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGT 600
QY 529 ATGGCTTGGGATATGATGATGAATCGTTCGCTACAAACGGCCCTGGTGGTATCCAGCTG 588
Db 601 ATGGCTTGGGATATGATGATGAATCGTTCGCTACAAACGGCCCTGGTGGTATCCAGCTG 660
QY 589 CTCGGATCCCAAGCTGCTGTGACATGTTGGGCGGGGCCCATTTGGGAGTCTCTGGCG 648
Db 661 CTCGGATCCCAAGCTGCTGTGACATGTTGGGCGGGGCCCATTTGGGAGTCTCTGGCG 720
QY 649 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTC 708
Db 721 GGTCTCGCTACTATTCCATGTTGGGAACTGGGCTAAGGTTTGAATGTGATGCTACTC 780
QY 709 TTTGCTCCCTAATAG 723
Db 781 TTTGCTCCCTAATAG 795
RESULT 5
ID AAL48939 standard; DNA; 2082 BP.
XX AAL48939;
XX 24-OCT-2002 (first entry)
DT Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.
DE Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX
KW

KW virucide; immunostimulant; vaccine; ds.
 XX Hepatitis C virus.
 OS WO20025548-A2.
 PN 18-JUL-2002.
 XX 11-JAN-2002; 2002WO-EP00219.
 PF 11-AUG-2001; 2001US-260699P.
 PR 30-AUG-2001; 2001US-315768P.
 XX (INNO-) INNOGENETICS NV.
 PA Maertens G, Bosman F, Buysee M;
 XX WPI; 2002-599657/64.
 DR P-PSDB; AAO18678.
 XX New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection
 XX Example 2; Page 206-209; 243pp; English.
 PS The present invention relates to new therapeutic vaccine compositions for
 XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;
 SQ

Query Match 86.4%; Score 624.8; DB 24; Length 2082;
 Best Local Similarity 90.5%; Pred. No. 1e-159;
 Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTTCATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGGTACATTCCG 63
 Db 4 TTGGGTAAGTTCATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGGTACATTCCG 63
 QY 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGCGTCCGGGTTCTG 123
 Db 64 CTCGTGGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGCGTCCGGGTTCTG 123
 QY 124 GAGGACGGGTGAATATGCAACAGGAATTTGCGCGGTGCTCTTTCTATCTTCCTC 183
 Db 124 GAGGACGGGTGAATATGCAACAGGAATTTGCGCGGTGCTCTTTCTATCTTCCTC 183
 QY 184 TTGGCTTTGCTGTCCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTTGCC 243
 Db 184 TTGGCTTTGCTGTCCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTTGCC 243
 QY 244 GGGATGTACATGTACAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCGGAC 303
 Db 244 GGGATGTACATGTACAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGGACGCGGAC 303
 QY 304 ATGATCATGCACACCCCGGGTGGTGGCTTCGCTTTCGGAGAACAACTCTTCCCGCTGC 363
 Db 304 ATGATCATGCACACCCCGGGTGGTGGCTTCGCTTTCGGAGAACAACTCTTCCCGCTGC 363
 QY 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAGCCAGGCTCCCGACGACAAATA 423
 Db 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAGCCAGGCTCCCGACGACAAATA 423

QY 424 CGACGCCACGTCGAT----- 438
 Db 424 CGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
 QY 439 -----TCCACAGCTGTTCCACCATCTCGCCTCGCGGCAAT 471
 Db 484 GACCTCTCGGGATCTGCTTCTCCGCTCCACCATCTCGCCTCGCGGCAAT 543
 QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGCCACATAACGGGTCAACGATG 531
 Db 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGCCACATAACGGGTCAACGATG 603
 QY 532 GCTTGGGATATGATGATGAACCTGTCGCTCAACGGCCCTGCTGATTCGAGCTGCTC 591
 Db 604 GCTTGGGATATGATGATGAACCTGTCGCTCAACGGCCCTGCTGATTCGAGCTGCTC 663
 QY 592 CGGATCCCAACAGCTGCTGTCGACATGTTGGGGGGGCCCATTTGGGAGTCTCTGGCGGT 651
 Db 664 CGGATCCCAACAGCTGCTGTCGACATGTTGGGGGGGCCCATTTGGGAGTCTCTGGCGGT 723
 QY 652 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 711
 Db 724 CTCGCTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 783
 QY 712 GC 713
 Db 784 GC 785

RESULT 6
 AAT12973
 ID AAT12973 standard; DNA; 2086 BP.
 XX AAT12973;
 DT 24-SEP-1996 (first entry)
 XX HCV E1 construct HCCI65.
 XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX Hepatitis C virus.
 OS WO9604385-A2.
 PN 15-FEB-1996.
 PD 31-JUL-1995; 95WO-EP03031.
 PF 29-JUL-1994; 94EP-0870132.
 PR (INNO-) INNOGENETICS NV.
 PA Bosman F, Buysee M, De Martynoff G, Maertens G;
 XX WPI; 1996-129401/13.
 DR Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 XX proteins - in presence of disulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX Claim 23; Fig 21; 146pp; English.
 XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a

disulphide bond cleavage agent, after lysis of recombinant host cells.
The constructs containing the purified HCV envelope proteins can be used
for vaccinating humans against HCV, for in vitro detection of HCV
antibodies in a sample, and in a serotyping assay for detecting one or
more serological types of HCV present in a biological sample. The
constructs can also be immobilised on a solid substrate and incorporated
into a reversed phase hybridisation assay for determining the presence or
the genotype of HCV. The new purification method preserves the
conformation of the recombinantly expressed E1, E2 and E1/E2, and
eliminates contaminating proteins. Antigens isolated using this method
are more reactive with human sera than those isolated by known
techniques.

XX Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;
SQ

Query Match 86.4%; Score 624.8; DB 17; Length 2086;
Best Local Similarity 90.5%; Pred. No. 1e-158; 2; Indels 72; Gaps 1;
Matches 708; Conservative 0; Mismatches 2;

4 TTGGTAAAGTCTATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGTACATTCCG 63
4 TTGGTAAAGTCTATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGTACATTCCG 63
64 CTGCTGGGCGCCCTAGGGCGCTGCGAGGCTCGGCGATGGCGTCCGGTCTG 123
64 CTGCTGGGCGCCCTAGGGCGCTGCGAGGCTCGGCGATGGCGTCCGGTCTG 123
124 GAGACGGCTGAACTATGCAACAGGAAATTCGCCGTTCTCTCTATCTCTCTC 183
124 GAGACGGCTGAACTATGCAACAGGAAATTCGCCGTTCTCTCTATCTCTCTC 183
184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTGTC 243
184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTGTC 243
244 GGGATGACCATGTCAGCAAGCTGCTCAACTCAAGCAATTCGTATGAGGAGCGGAC 303
244 GGGATGACCATGTCAGCAAGCTGCTCAACTCAAGCAATTCGTATGAGGAGCGGAC 303
304 ATGATCATGACACCCCGGCTGGTCCCTGCTTCGGGAGAACACTCTTCGCGTGC 363
304 ATGATCATGACACCCCGGCTGGTCCCTGCTTCGGGAGAACACTCTTCGCGTGC 363
364 TGGGTAGGCTCACCCCAACGCTCGAGTAGGAAACGCGAGCTCCCAACGAGCAATA 423
364 TGGGTAGGCTCACCCCAACGCTCGAGTAGGAAACGCGAGCTCCCAACGAGCAATA 423
424 CGAGCCAGCTCGAT----- 438
424 CGAGCCAGCTCGATTTGCTGGGCGGCTGCTTCTGCTCGCTATGATCGTGGG 483
439 -----TCCAGAGTGTTCACCATCTGCTCGCTCGCGGCAT 471
484 GACCTCTGGGATCTGCTTCTCTCTCTCCAGCTGTTTCCACCATCTGCTCGCGGCAT 543
472 GAGACGGTCCAGGCTGCAATGCTCAATCTATCCCGGCAATTAACGGGTACCGTATG 531
544 GAGACGGTCCAGGCTGCAATGCTCAATCTATCCCGGCAATTAACGGGTACCGTATG 603
532 GCTTGGGATATGATGATGAATGCTGCGCTACAAACGGGCTGGTATCGAGTGTCTC 591
604 GCTTGGGATATGATGATGAATGCTGCGCTACAAACGGGCTGGTATCGAGTGTCTC 663
592 CGGATCCCAACAGCTCTCGTGGACATGTTGGGCGGCGCCATTGGGGAGTCTCGCGGGT 651
664 CGGATCCCAACAGCTCTCGTGGACATGTTGGGCGGCGCCATTGGGGAGTCTCGCGGGC 723
652 CTGCGCTACTATTCCATGTTGGGAACTGCGGCTAAGTTTTGTTGTGATGCTACTCTTT 711
724 CTGCGCTACTATTCCATGTTGGGAACTGCGGCTAAGTTTTGTTGTGATGCTACTCTTT 783
712 GC 713
||

784 GC 785

RESULT 7
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX
AC AAT12974;
XX
DT 25-SBP-1996 (first entry)
XX
DE HCV E1 construct HCC166.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
FN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di-sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by a
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
SQ

Query Match 86.4%; Score 624.8; DB 17; Length 2433;
Best Local Similarity 90.5%; Pred. No. 1.1e-158;
Matches 708; Conservative 0; Mismatches 2; Indels 72; Gaps 1;

4 TTGGTAAAGTCTATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGTACATTCCG 63
355 TTGGTAAAGTCTATCGATACCCCTTACATCGGCTTCGCGACCTCGTGGGTACATTCCG 414
64 CTGCTGGGCGCCCTAGGGCGCTGCGAGGCGCTTCAGGCGCTTCGCGTCCGGTCTG 123
415 CTCGTGGGCGCCCTAGGGCGCTGCGAGGCGCTTCAGGCGCTTCGCGTCCGGTCTG 474

```
QY 124 GAGGACGGTGAACATGATGACAGGAAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 183
Db 475 GAGGACGGGTAACATGATGACAGGAAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 534
QY 184 TTGGCTTTTCTGCTGCTGCTGACCGTTCCAGCTTTCAGCTTATGAAGTGGCAACGTGTC 243
Db 535 TTGGCTTTTCTGCTGCTGCTGACCGTTCCAGCTTTCAGCTTATGAAGTGGCAACGTGTC 594
QY 244 GGGATGATCATGTGATCAAGAACGACTGCTCAACTCAAGCAATGATGATGAGGACGGAC 303
Db 595 GGGATGATCATGTGATCAAGAACGACTGCTCAACTCAAGCAATGATGATGAGGACGGAC 654
QY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 655 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
QY 364 TGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAAGCCAGCGTCCCGCCACGACAATA 423
Db 715 TGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAAGCCAGCGTCCCGCCACGACAATA 774
QY 424 CGACGCCACGTCGAT----- 438
Db 775 CGACGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCGGCAAT 471
Db 835 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
QY 472 GAGACGGTGCAGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATG 931
Db 895 GAGACGGTGCAGACTGCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATG 954
QY 532 GTTGGATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 991
Db 955 GTTGGATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
QY 592 CGGATCCCAAGCTGCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 651
Db 1015 CGGATCCCAAGCTGCTGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1074
QY 652 CTCGCTTACTTATTCATGTTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 711
Db 1075 CTCGCTTACTTATTCATGTTGGGAACTGGGCTAAGGTTTGTATGATGCTACTCTTT 711
QY 712 GC 713
Db 1135 GC 1136

RESULT 8
ID AAL48940
XX AAL48940 standard; DNA; 2434 BP.
AC AAL48940;
XX
DT 24-OCT-2002 (first entry)
DE
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
XX
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
XX WO200255548-A2.
XX
XX 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
XX
XX 30-AUG-2001; 2001US-315768P.
XX
```

```
PA (INNO-) INNOGENETICS NV.
XX Maertens G, Bosman F, Buyse M;
XX WPI: 2002-599657/64.
DR P-PSDB; AAO18679.
XX
PT New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX
PS Example 2; Page 212-215; 243pp; English.
XX
CC The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunizing humans
CC against HCV. The recombinants, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;

Query Match 84.9%; Score 613.8; DB 24; Length 2434;
Best Local Similarity 90.4%; Pred. No. 1e-155; 2; Indels 73; Gaps 2;
Matches 708; Conservative 0; Mismatches 0;

QY 4 TTGGGTAAGGTATCATGATACCCCTTACATCGGGCTTCGCGACCTCGTGGGTACATTCCG 63
Db 355 TTGGGTAAGGTATCATGATACCCCTTACATCGGGCTTCGCGACCTCGTGGGTACATTCCG 414
QY 64 CTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGCGCATGCGGCTCGGGTTCG 123
Db 415 CTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGCGCATGCGGCTCGGGTTCG 474
QY 124 GAGGACGGGCTGAACATGATGCAACAGGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 183
Db 475 GAGGACGGGCTGAACATGATGCAACAGGGAATTTGCCCGGTTGCTTTCTCTATCTTCCTC 534
QY 184 TTGGCTTTTCTGCTGCTGCTGACCGTTTCAGCTTTCAGCTTATGAAGTGGCAACGTGTC 242
Db 535 TTGGCTTTTCTGCTGCTGCTGACCGTTTCAGCTTTCAGCTTATGAAGTGGCAACGTGTC 594
QY 243 CGGGATGTATACCATGTGTCAGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGA 302
Db 595 CGGGATGTATACCATGTGTCAGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGA 654
QY 303 CATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362
Db 655 CATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
QY 363 CTGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAAGCCAGCGTCCCGCCACGACAAT 422
Db 715 CTGGGTAGCGCTCACCCCGGCTGCGAGCTAGGAAGCCAGCGTCCCGCCACGACAAT 774
QY 423 ACGACGCCACGTCGAT----- 438
Db 775 ACGACGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCGGCAAT 470
Db 835 GAGCTCTGCGGATCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
QY 471 TGAGACGGTGCAGGACTGCAATGCTCAATCTATCCCGGCCACATACCGGTCACCGTAT 930
Db 895 TGAGACGGTGCAGGACTGCAATGCTCAATCTATCCCGGCCACATACCGGTCACCGTAT 954
```

Seq	Sequence	606 BP	109 A	193 C	167 G	137 T	0 Other
	Query Match	82.7%	Score 598.2	DB 17	Length 606		
	Best Local Similarity	99.5%	Pred. No. 1.1e-151				
	Matches 600	Conservative	0	Mismatches 3	Indels 0	Gaps 0	
QY	1	ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT	60				
DB	1	ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGGTACATT	60				
QY	61	CGGCTCGTCCGGCGCCCGCCCTAGGGGGCGCTGCGAGGGCCCTGGCGATGGCGTCCGGGTT	120				
DB	61	CGGCTCGTCCGGCGCCCGCCCTAGGGGGCGCTGCGAGGGCCCTGGCGATGGCGTCCGGGTT	120				
QY	121	CTGGAGGACGGCGTGAACTATCGAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTC	180				
DB	121	CTGGAGGACGGCGTGAACTATCGAACAGGGAATTTGCCCGGTTCCTTTCTCTATCTTC	180				
QY	181	CTCTTCGGCTTTTGCTGCTCTGTCTGACCGTTCACGCTTCGCGTTATGAAGTGCACAACGTG	240				
DB	181	CTCTTCGGCTTTTGCTGCTCTGTCTGACCGTTCACGCTTCGCGTTATGAAGTGCACAACGTG	240				
QY	241	TCCGGGATGTACATGTACGAAAGCACTGCTCAACTCAAGCAATTGTATGAGGCAGCG	300				
DB	241	TCCGGGATGTACATGTACGAAAGCACTGCTCAACTCAAGCAATTGTATGAGGCAGCG	300				
QY	301	GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC	360				
DB	301	GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTCTTCCCGC	360				
QY	361	TGCTGGGTAGCGTTCACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCCACCAACACA	420				
DB	361	TGCTGGGTAGCGTTCACCCCGACGCTCGAGCTAGGAACGCCAGCGTCCCCACCAACACA	420				
QY	421	ATACGAGCCACGTCGATTCCAGCTGTTCAACATCTCGCTCGCCGGCATGAGACGGTG	480				
DB	421	ATACGAGCCACGTCGATTCCAGCTGTTCAACATCTCGCTCGCCGGCATGAGACGGTG	480				
QY	481	CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540				
DB	481	CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540				
QY	541	ATGATGATGAACTGGTGCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTCCGATCCCA	600				
DB	541	ATGATGATGAACTGGTGCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTCCGATCCCA	600				
QY	601	CAA	603				
DB	601	TAA	603				

RESULT 10	
AAL48927	Hepatitis C virus clone HCC139 E1 protein coding sequence.
ID AAL48927	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX AC	virucide; immunostimulant; vaccine; ds.
XX AC	
XX DT	Hepatitis C virus.
XX DT	Hepatitis C virus.
XX OS	
XX PN	WO200255548-A2.
XX PD	
XX PF	18-JUL-2002.
XX PP	
XX PR	11-JAN-2002; 2002WO-EP00219.
XX PR	
XX PR	11-JAN-2001; 2001US-260699P.
XX PR	

531	GCCTTGGGATATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCT	590
QY		
595	GCCTTGGGATATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCT	1014
Db		
591	CGGGATCCACAAAGCTGTCGTGGACATGTTGGCGGGGGCCCATTTGGGGAGTCTTGGCGGG	650
QY		
1015	CGGGATCCACAAAGCTGTCGTGGACATGTTGGCGGGGGCCCATTTGGGGAGTCTTGGCGGG	1074
Db		
651	TCCTCGCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTTCGATTGTGATGCTACTCTT	710
QY		
1075	CCTCGCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTTCGATTGTGATGCTACTCTT	1134
Db		
711	TGC 713	
QY		
1135	TGC 1137	
Db		

RESULT 9	
AAT12963	
ID	AAT12963 standard; DNA; 606 BP.
XX	
AC	AAT12963;
XX	
DT	24-SEP-1996 (first entry)
XX	
DE	HCV E1 construct HCC139.
XX	
XX	HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW	serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW	ss.
XX	
OS	Hepatitis C virus.
XX	
PN	WO9604385-A2.
XX	
PD	15-FEB-1996.
XX	
XX	31-JUL-1995; 95WO-EP03031.
PF	
XX	
XX	29-JUL-1994; 94EP-0870132.
PR	
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Bosman F, Buyse M, De Martynoff G, Maertens G;
XX	
PR	WPI. 1996-129401/13.

XX
PS
CC Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for *in vitro* detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

Mon Dec 22 13:28:37 2003

```
PS 11 CGCTCTGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCTCCGGGTT 120
XX |||
CC |||
CC CGCTCTGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCTCCGGGTT 120
CC |||
CC 121 CTGGAGGAGCGGCTGAACCTATGACACAGGGAATTTGGCCGGTTCCTCTTCTATCTTC 180
CC |||
CC 121 CTGGAGGAGCGGCTGAACCTATGACACAGGGAATTTGGCCGGTTCCTCTTCTATCTTC 180
CC |||
CC 181 CTCCTTGGCTTCTGTCCTGTCGACCGTTCCAGCTTCGGTTATGAAGTCGCAACGTG 240
CC |||
CC 181 CTCCTTGGCTTCTGTCCTGTCGACCGTTCCAGCTTCGGTTATGAAGTCGCAACGTG 240
CC |||
CC 241 TCCGGGATGATACATGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGACG 300
CC |||
CC 241 TCCGGGATGATACATGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGACG 300
CC |||
CC 301 GACATGATCATGACACCCCGGGTGCCTGCTGGTTCGGGAGAACATCTTCCCGC 360
CC |||
CC 301 GACATGATCATGACACCCCGGGTGCCTGCTGGTTCGGGAGAACATCTTCCCGC 360
CC |||
CC 361 TGTGAGGAGCGCTACCCCGAGCTCGAGTGAAGACCGAGCTGCCACACGACG 420
CC |||
CC 361 TGTGAGGAGCGCTACCCCGAGCTCGAGTGAAGACCGAGCTGCCACACGAG 420
CC |||
CC 421 ATAGGAGCGGACGTCGATTCACAGCTGTTTCAACCATCTCGCTCGCGGATGAGAGCGTG 480
CC |||
CC 421 ATAGGAGCGGACGTCGATTCACAGCTGTTTCAACCATCTCGCTCGCGGATGAGAGCGTG 480
CC |||
CC 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATAAACGGGTGACCGTATGCTTGGGAT 540
CC |||
CC 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATAAACGGGTGACCGTATGCTTGGGAT 540
CC |||
CC 541 ATGATGATGAGTGGTGCCTTACAAACGGCCCTGGTGTATCGCAGTGTCTCCGGATC 597
CC |||
CC 541 ATGATGATGAGTGGTGCCTTACAAACGGCCCTGGTGTATCGCAGTGTCTCCGGATC 597
CC |||
```

```
RESULT 12
AAL48928
ID AAL48928 standard; DNA; 636 BP.
XX
AC AAL48928;
XX
XX 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus clone HCCI40 E1 protein coding sequence.
XX
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX
XX virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
XX WO200255548-A2.
XX
XX 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
XX
XX 30-AUG-2001; 2001US-315768P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buysee M;
XX
XX WPI; 2002-599657/64.
XX
XX P-PSDB; AAO18669.
XX
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection
XX
```

```
PS Example 2; Page 179-180; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV) -specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX in vitro monitoring of HCV disease or prognosing the response to
XX treatment of patients suffering from HCV infection. The present sequence
XX is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;
XX
XX Query Match 82.6%; Score 597; DB 24; Length 636;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-151;
XX Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
XX |||
XX DB 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
XX |||
XX QY 61 CCGCTCTGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCTCCGGGTT 120
XX |||
XX DB 61 CCGCTCTGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCTCCGGGTT 120
XX |||
XX QY 121 CTGAGGAGCGGCTGAACCTATGACACAGGGAATTTGGCCGGTTCCTCTTCTATCTTC 180
XX |||
XX DB 121 CTGAGGAGCGGCTGAACCTATGACACAGGGAATTTGGCCGGTTCCTCTTCTATCTTC 180
XX |||
XX QY 181 CTCCTTGGCTTCTGTCCTGTCGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTG 240
XX |||
XX DB 181 CTCCTTGGCTTCTGTCCTGTCGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTG 240
XX |||
XX QY 241 TCCGGGATGATACATGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGACG 300
XX |||
XX DB 241 TCCGGGATGATACATGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGACG 300
XX |||
XX QY 301 GACATGATCATGACACCCCGGGTGCCTGCTGGTTCGGGAGAACAACTTTCGGC 360
XX |||
XX DB 301 GACATGATCATGACACCCCGGGTGCCTGCTGGTTCGGGAGAACAACTTTCGGC 360
XX |||
XX QY 361 TGTGAGGAGCGCTACCCCGAGCTCGAGTGAAGACCGAGCTGCCACACGACG 420
XX |||
XX DB 361 TGTGAGGAGCGCTACCCCGAGCTCGAGTGAAGACCGAGCTGCCACACGACG 420
XX |||
XX QY 421 ATAGGAGCGGACGTCGATTCACAGCTGTTTCAACCATCTCGCTCGCGGATGAGAGCGTG 480
XX |||
XX DB 421 ATAGGAGCGGACGTCGATTCACAGCTGTTTCAACCATCTCGCTCGCGGATGAGAGCGTG 480
XX |||
XX QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATAAACGGGTGACCGTATGCTTGGGAT 540
XX |||
XX DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATAAACGGGTGACCGTATGCTTGGGAT 540
XX |||
XX QY 541 ATGATGATGAGTGGTGCCTTACAAACGGCCCTGGTGTATCGCAGTGTCTCCGGATC 597
XX |||
XX DB 541 ATGATGATGAGTGGTGCCTTACAAACGGCCCTGGTGTATCGCAGTGTCTCCGGATC 597
XX |||
```

```
RESULT 13
ABK91411
ID ABK91411 standard; DNA; 9605 BP.
XX
XX AC ABK91411;
XX
XX DT 15-NOV-2002 (first entry)
XX
XX DE Hepatitis C virus Con 1 isolate DNA.
XX
XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
XX
```


hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV; internal ribosome entry site; IRES; NS5A; HCV replication.

Hepatitis C virus.

Key

CDS 342..9374

```
/*tag= a
/product=
```

[illegible]

W0200259321-22

[illegible]

01-AUG-2002.

16-JAN-2002; 20

23-JAN-2001;

(RICE-) IST RICERCHE BIOL MO

De Francesco R, Migliaccio

WPI: 2002-599793/64.

P-PSDB; ABG32451.

New Hepatitis C virus

NS5 encoding region
ribosome entry site

and expression -
Claim 9; Page 36-39; 69pp; English.

The invention relates to nucleic acid molecules comprising altered HCV

NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region coding for one or more NS3,

NS5A, or EMCV IRE. mutations are det

(1) an expression vector comprising

the altered nucleic acids, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is the HCV replicon Con 1, used as a basis for the adaptive mutations of the invention.

Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 other;

XX	01-AUG-2002.
PD	
XX	16-JAN-2002; 2002WO-EP00526.
PF	
XX	23-JAN-2001; 2001IUS-263479P.
PR	
XX	(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
PA	
XX	De Francesco R, Migliaccio G, Paonessa G;
PI	
XX	WPI; 2002-599793/84.
DR	
XX	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT	ribosome entry site (IRES) region, useful in studying HCV replication
PT	and expression -
XX	
PS	Claim 9; Page -: 69pp; English.
XX	The invention relates to nucleic acid molecules comprising altered HCV
CC	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC	internal ribosome entry site (IRES) region coding for one or more NS3,
CC	NS5A, or EMCV IRES mutations, respectively. The location of the
CC	mutations are detailed in the specification. Also included are
CC	(1) an expression vector comprising a nucleotide sequence coding for
CC	the altered nucleic acids, which is transcriptionally coupled to an
CC	exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC	the altered nucleic acids; (3) a recombinant cell produced by introducing
CC	into a human hepatoma cell the altered nucleic acids; (4) producing a
CC	HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC	functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC	method; and (6) measuring the ability of a compound to affect HCV
CC	activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC	studying HCV replication and expression, and HCV and host cell
CC	interactions, producing HCV RNA and proteins, and providing a system
CC	for measuring the ability of a compound to modulate one or more HCV
CC	activities e.g. to discover drugs which may treat HCV mediated
CC	diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC	The present sequence is an HCV replicon Con 1 mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the HCV sequence appearing as
CC	ABK91411 and the information in Claim 9.
XX	
SQ	Sequence 9605 BP; 1910 A; 2884 C; 2732 G; 2079 T; 0 other;
Query Match 77.3%; Score 559.2; DB 24; Length 9605;	
Best Local Similarity 85.3%; Pred. No. 8.9e-141;	
Matches 667; Conservative 0; Mismatches 43; Indels 72; Gaps 1;	
QY	4 TTGGGTAGGTCATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGGTACATTCCG 63
Dd	696 TTGGGTAGGTCATCGATACCTTACATCGGGCTTCGCCGACTCATGGGGTACATTCCG 755
QY	64 CTCGTGGCGCCCCCTCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGCGTCCGGTTCG 123
Dd	756 CTCGTGGCGCCCCCTCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGCGTCCGGTTCG 815
QY	124 GAGGACGGCGTGAACTATGCAACAGGGAATTGCCCGGTTCCTCTTTCTATCTTCCTC 183
Dd	816 GAGGACGGCGTGAACTATGCAACAGGGAATTGCCCGGTTCCTCTTTCTATCTTCCTC 875
QY	184 TTGGCTTTGCTGTCCTGTCTGACCGTTTCAGCTTCGGCTTATGAAGTCGCAACGTCTCC 243
Dd	876 TTGGCTTTGCTGTCCTGTGTGACCATCCAGCTTCGGCTTATGAAGTCGCAACGTATCC 935
QY	244 GGAGTGTACATGTCACGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGCGCGGAC 303
Dd	936 GGAGTGTACATGTCACGAACGACTGCTCCAACGCACTGCTTGTATGAGCGCGGAC 995
QY	304 ATGATCATGCACACCCCGGGTCGTCCTCGGTTCGGGAGAACCACTCTTCCCGTTCG 363
Dd	996 ATGATCATGCATACCCCGGGTCGTCCTCGGTTCGGGAGAACCACTCTTCCCGTTCG 1055

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication
PT and expression -
XX
PS Claim 9; Page -; 69pp; English.
XX
CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the
CC mutations are detailed in the specification. Also included are
CC (1) an expression vector comprising a nucleotide sequence coding for
CC the altered nucleic acids, which is transcriptionally coupled to an
CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
CC the altered nucleic acids; (3) a recombinant cell produced by introducing
CC into a human hepatoma cell the altered nucleic acids; (4) producing an
CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
CC method; and (6) measuring the ability of a compound to affect HCV
CC activity. The HCV replicons and HCV replicon enhanced cells are useful in
CC studying HCV replication and expression, and HCV and host cell
CC interactions, producing HCV RNA and proteins, and providing a system
CC for measuring the ability of a compound to modulate one or more HCV
CC activities e.g. to discover drugs which may treat HCV mediated
CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
CC The present sequence is an HCV replicon Con 1 mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the HCV sequence appearing as
CC ABK91411 and the information in Claim 9.
XX
SQ Sequence 9605 BP; 1909 A; 2883 C; 2734 G; 2079 T; 0 other;

Query Match 77.3%; Score 559.2; DB 24; Length 9605;
Best Local Similarity 85.3%; Pred. No. 8.9e-141;
Matches 667; Conservative 0; Mismatches 43; Indels 72; Gaps 1;

QY 4 TTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCGAACCTCGTGGGGTACATTCGG 63
DB TTTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCGAACCTCGTGGGGTACATTCGG 755
QY 64 CTCGTGGCGCCCTAGGGGGCGTGGCGAGGCGCTGGCGCATGGCTCGGGTTCTG 123
DB TTTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCGAACCTCGTGGGGTACATTCGG 815
QY 124 GAGGACGGCGTGAACCTATGCAACAGGGAATTTGGCGGTTGCTTCTCTATCTTCCTC 183
DB GAGGACGGCGTGAACCTATGCAACAGGGAATTTGGCGGTTGCTTCTCTATCTTCCTC 875
QY 184 TTGGCTTTTGTCTGTCTGTCTGACCGTTCCAGCTTCGCGTTATGAAGTGGCAACGTGTC 243
DB TTGGCTTTTGTCTGTCTGTCTGACCGTTCCAGCTTCGCGTTATGAAGTGGCAACGTATCC 935
QY 244 GGGATGTACATGTACGAAAGCACTGCTCCAACTCAAGCAATTGTATGAGGAGCGGAC 303
DB GGAGTGTACATGTACGAAAGCACTGCTCCAACTCAAGCAATTGTATGAGGAGCGGAC 995
QY 304 ATGATCATGCACACCCCGGGTGGCGTGGCTGGGAGAACAACTCTTCCCGCTGC 363
DB ATGATCATGCATACCCCGGGTGGCGTGGCTGGGAGAACAACTCTTCCCGCTGC 1055
QY 364 TGGGTAGCGCTCAACCCCGGCTAGGAGCCAGCGTCCCGACCAACCACTCTTCCCGCTGC 423
DB TGGGTAGCGCTCACTCCAGCGTCCGCGGCGGAGAACGCTAGCTCCCGCTAGCAGATA 1115
QY 424 CGACGCCACGTGATTT----- 439
DB CGACGCCACGTGATTTGCTGTTGGGGGGGCTGCTCTCTGCTCCGCTATGTACGTGGGA 1175
QY 440 -----CCAGCTGTTTACCATCTCGCTCGCGCGGAT 471
DB GATCTCTCGGGATCTGTTTTTCTCTGCTGCGGCCAGCTGTTTACCTTCTCGCTCGCGGCCAC 1235

QY 472 GAGACGGTCCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACCGTATG 531
DB GAGACAGTACAGGACTGCAATTGCTCAATATATCCGGCCACGTGACAGGTCAACCGTATG 1295
QY 532 GCTTGGGATATGATGAACTGCTCGCTCAACCGGCCCTGGTGGTATCGCAGCTGCTC 591
DB GCTTGGGATATGATGAACTGCTCGCTCAACCGGCCCTAGTGGTATCGCAGTTACTC 1355
QY 592 CGGATCCCAACAAGCTGTCTGGGACATGGTGGGGGGGCCCATTTGGGGAGTCTCTGGCGGGT 651
DB CGGATCCCAACAAGCTGTCTGGGATATGGTGGGGGGGCCCATTTGGGGAGTCTCTAGCGGGC 1415
QY 652 CTCGCTTACTATTCCATGGTGGGAACTGGGCTAAGGTTTTCATTGTGATGCTACTCTTT 711
DB CTTGCTTACTATTCCATGGTGGGAACTGGGCTAAGGTTTCTGATTGTGATGCTACTCTTT 1475
QY 712 GC 713
DB 1476 GC 1477

Search completed: December 19, 2003, 18:51:12
Job time : 203.88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1859.96 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-21

Perfect score: 723

Sequence: 1 ATGTTGGTAAGTATCGA.....TACTCTTCTGCTCCCTAATAG 723

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101.2	14.0	488	9	AV755731
C 2	79.6	11.0	492	9	AV758366
C 3	41.6	5.8	502	12	BI879124
C 4	40.6	5.6	275	9	AV835132

5	40.6	5.6	402	9	AV392783
6	40.6	5.6	551	9	AV392165
7	40.6	5.6	552	12	BI996341
8	40.6	5.6	584	12	BI727879
9	40.4	5.6	1201	13	BX356664
10	40.2	5.6	1162	12	BM918259
11	40.2	5.5	1201	9	AL513886
C 11	39	5.4	359	12	BJ252669
C 12	39	5.4	375	12	BJ246716
C 13	39	5.4	840	29	CC335916
C 14	39	5.4	873	14	CD446071
C 15	39	5.4	873	14	CD446071
C 16	38.6	5.3	925	29	CNS0091P
C 17	38.4	5.3	925	29	CNS0091P
C 18	38.4	5.3	702	14	CD432549
C 19	38.4	5.3	740	12	BJ536071
C 20	38.4	5.3	925	29	CNS0091P
C 21	38.4	5.3	970	29	CNS010C9
C 22	38.4	5.3	987	29	CNS015VX
C 23	38.2	5.3	533	6	AU192776
C 24	38.2	5.3	538	6	AU193705
C 25	38.2	5.3	544	6	AU190971
C 26	38.2	5.3	544	6	AU192419
C 27	38.2	5.3	1270	12	BG968359
C 28	38	5.3	354	14	CB966525
C 29	38	5.3	742	13	BQ752873
C 30	38	5.3	1201	13	BX381961
C 31	37.8	5.2	435	14	C72860
C 32	37.8	5.2	533	29	CC010084
C 33	37.8	5.2	533	29	CC405164
C 34	37.8	5.2	826	29	BZ736582
C 35	37.8	5.2	895	29	CC359028
C 36	37.8	5.2	925	29	CC359026
C 37	37.8	5.2	940	29	CC010085
C 38	37.8	5.2	951	29	CC405167
C 39	37.6	5.2	431	9	AV639153
C 40	37.4	5.2	360	9	AJ473805
C 41	37.4	5.2	637	13	BQ293470
C 42	37.4	5.2	641	13	BQ172543
C 43	37.4	5.2	650	14	CA828039
C 44	37.4	5.2	834	29	BZ641450
C 45	37.4	5.2	841	29	BZ641457

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfliA; Site_2: sfliB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN
Query Match 14.0%; Score 101.2; DB 9; Length 488;
Best Local Similarity 65.9%; Pred. No. 2.7e-15; Indels 1; Gaps 1;
Matches 162; Conservative 0; Mismatches 83;
QY 445 CTGTTACCAATCTCGCTCGCGGATGAGACGGTGCAGGACTGCAATTTGCTCAATCTAT 504
DB 403 CAGCTGATCATCTGGCTCAGACCAATGAGTTTGTGCATGAATGCAACTGCTCCATCTAT 344
QY 505 CCCGGCCACATACCGGTACCGTATG-GCTGGGATATGATGATGAACTGCTGCGCTAC 563
DB 343 CTTGGGCCATCATCTGACACCGTATGAGCATGGGACATGATGAATGCTGTGTCAC 284
QY 564 AACGGCCCTGCTGATTCGAGCTGCTCGGATCCCAAGCTGCTGCTGGACATGCTGGC 623
DB 283 CGCTGCTATGATCATGCGGTACGCAATGCGGTTCTGAGGTCTCATATAGATATCATCAG 224
QY 624 GGGGGCCCATGGGGAGTCTGGGGGTCTCGGCTACTTCCATTCATGTTGGGAACCTGGC 683
DB 223 CGGGGCTCACTGGGGCGTATGTTGGGCTTAGCGTACTTCTATGACGAGGAGCTGGGC 164
QY 684 TAAGCT 689
DB 163 GAAAGT 158

RESULT 2
AV758366/c
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zegu Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. .492
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKB03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"

/clone_lib="BM"
/note="Vector: pTriplex2; Site_1: sfliA; Site_2: sfliB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN
Query Match 11.0%; Score 79.6; DB 9; Length 492;
Best Local Similarity 60.7%; Pred. No. 8.8e-10; Indels 1; Gaps 1;
Matches 147; Conservative 0; Mismatches 94;
QY 449 TCACCATCTCGCTCGCGGATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCG 508
DB 400 TGATTATCTCTCAGCAGCAACATTGGTTTGTGCAAGATGCAACTGCTCATTTCTATCTG 341
QY 509 GCCACATAACGGTTC-ACCGTATGGTGGGATATGATGAACTGCTGCGCTCAACAACG 567
DB 340 GCTGCATCATCTGCACTACAGTATGGATAGGCTATGATGATGAAGTGGTGGCAGCGGT 281
QY 568 GCCCTGCTGCTATCGCAGCTGCTCGGATCCCAAGCTGCTGTCGACATGCTGGCGGG 627
DB 280 TCCATGATCTGGCGTACGCAATGCGGTTCTTGAAGTCTCATAGATATCAAGCTGG 221
QY 628 GCCCATTTGGGAGTCTCGGGGCTCTCGCTACTTATTCATGTTGGGAACCTGGGCTAAG 687
DB 220 GCACACTGCGGCGTATGTTTGGCTCAGCTTACTTCAATGACGAGCGCTGGCCAAA 161
QY 688 GT 689
DB 160 GT 159

RESULT 3
BI879124/c
LOCUS BI879124 Zebrafish adult retina cDNA Danio rerio cDNA clone
DEFINITION IMAGE:4145367 5' similar to TR:Q9PMN4 Q9PMN4 RHODOPSIN. ; mRNA
sequence.
ACCESSION BI879124
VERSION BI879124.1 GI:16086395
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 502)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbratfish@wustl.edu
Library constructed by: Chandra Tucker and Gregory Niemi
Sequencing by: Washington University Genome Sequencing Center
Distribution: RessourcenZentrumPrimarDatenbank, Berlin, Germany
(web address: www.rzpd.de)
Trace considered overall poor quality
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1. .502
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/strain="wild-type"
/db_xref="taxon:7955"
/clone="IMAGE:4145367"

```

```

/sex="mixed"
/dev_stage="1-2 years"
/lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF)'"
/clone_lib="Zebrafish adult retina cDNA"
/notes="Vector: Lambda ZAP II (pBluescript SK-); Site_1:
EcoRI; Site_2: SalI; This Zebrafish library was
constructed by Dr. Susan B. Brockerhoff (email:
sbrock@u.washington.edu) RZPD library number: 760"
BASE COUNT      98 a 163 c 125 g 116 t
ORIGIN

Query Match      5.8%; Score 41.6; DB 12; Length 502;
Best Local Similarity 51.0%; Pred. No. 4.4;
Matches 98; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 300 GGACATGATCATCACACCCCGGGTGGCGTCCCTCGTTCGGGAGAACAACTCTTCCCG 359
Db 484 GGCCATGAGATGTGAAGCCACCATGATGGATGTTTCGCCAGCAGAAAGTTGTTG 425
QY 360 CTGCTGGGTAGGCTCACCCACGCTCGGAGTAGAAGCCAGGTCCTCCACACGAC 419
Db 424 CCAGTTTGAGTCCACCATCCAGCGCTCAATGGCATGTATCGAGCGACCATAGCCCAT 365
QY 420 AATACGACCCAGCTGATTCCTCAGCTGTTCCACATCTCGCTCGCGCATGAGCGGT 479
Db 364 CTCACCGCAGGTGGCTAAGTAGCTTACAGTTGCAGTTCAGTTCGGCAGAACAGAA 305
QY 480 GCAGGACTGCAA 491
Db 304 GTAGCCGTGCAA 293

RESULT 4
AV835132
LOCUS
DEFINITION
AV835132 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
spontaneum top three leaves adult, heading stage Hordeum vulgare
subsp. spontaneum cDNA clone bah24018, mRNA sequence.
ACCESSION
AV835132
VERSION
AV835132.1 GI:14527221
KEYWORDS
EST.
SOURCE
Hordeum vulgare subsp. spontaneum
ORGANISM
Hordeum vulgare subsp. spontaneum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 275)
REFERENCE
Sato, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished
Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazsato@rib.okayama-u.ac.jp,
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct
submission;
database: http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES
source
Location/Qualifiers
1..275
/organism="Hordeum vulgare subsp. spontaneum"
/mol_type="mRNA"
/cultivar="H602"
/db_xref="taxon:77009"
/clone="bah24018"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. spontaneum top three leaves adult, heading
stage"
30 a 123 c 67 g 46 t 9 others

BASE COUNT
ORIGIN

Query Match      5.6%; Score 40.6; DB 9; Length 275;
Best Local Similarity 48.2%; Pred. No. 6.6;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 306 GATCATGCACACCCCGGGTGGCGTCCCTCGTTCGGGAGAACAACTCTTCCCGTCTG 365
Db 1 GGTCTCTGAGCGNACCCGCCCTCTGCTCCGTTGCTCTCCGCTGCGCGACCGCG 60
QY 366 GGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCGAGCGTCCCGACACGAAATACG 425
Db 61 CGGCCAGCGGCACTCACCTNCTCTGCTCGTTGCGCCCTCCCTCTCGCGCCCCCTCG 120
QY 426 AGCCACAGTGTATCCAGCTGTTTACCATCTCGCTCGCGGCGATGACAGGTGAGGA 485
Db 121 ACGGCTGGGCTAGCTGGGTGAGTTACATCTCTGCTCTCGGCGACGAGCGTGGCGGA 180
QY 486 CTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCAACCGTATG 531
Db 181 ACCGNACTCGCTCCCTCTCCGCCCCCAGCGCTCGCACACGGAACG 226

RESULT 5
AV392783
LOCUS
DEFINITION
AV392783 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM09604_r 5', mRNA sequence.
ACCESSION
AV392783
VERSION
AV392783.1 GI:6546999
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii
ORGANISM
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 402)
REFERENCE
Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
DNA Res. 6 (6), 369-373 (1999)
JOURNAL
MEDLINE
20152988
PUBMED
10691129
COMMENT
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
Location/Qualifiers
1..402
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM09604_r"
/dev_stage="Photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      66 a 126 c 146 g 64 t
ORIGIN

Query Match      5.6%; Score 40.6; DB 9; Length 402;
Best Local Similarity 45.3%; Pred. No. 7.5;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTATCATTCGCTCTGCGCGCCCCCTTAGGGGGCGCTGCCAGGCGCT 101
Db 53 CGAGCTCATCTCGTTCATTTGTGCGCGCACTGCCAACATGAAGGACGTGCTGACGACCT 112
QY 102 GCGCATGGCTCCGGTTCGGAGNCGCGGTGAACATATGCAACAGGGAATTTGCCGG 161
Db 113 GGCCGCGCGCGCGAGTGGAGGCGCGCTACGCCACGAGTCCGTGAGTTGGCGC 172

```

us-09-899-303a-21.rst

Mon Dec 22 13:28:39 2003

```

162 TTGCTCTTCTATCTCTCTTCTCTTGGCTTGGTGTCTGTCTGTACCGTTCCAGCTTCGGC 221
173 CGGCAAGGTGTTTGAAGAGTCAAGAGTACGTTGTGAACCTCAGGCCAGAACCCAG 232
222 TTATGAAGTGCACAGCTGTCGGGATGTACCATGTACGAACGATGCTCCAACTCAAG 281
233 CTTGCGCGTCCGCTGGTGGCCACTCGCTGGGGCGGCGACCGCGGCTGCTGTGAT 292
282 CATTTGTATGAGGACGCGGATGATCATGTACACACCCCGGTCGTCGCTGGTTTCG 341
293 CTTGATGACACGACGAGGAGTTTGGCGGCGGATCTACGGCGGCGTGCCTCCAGTCCCGG 352
342 GGAGAACAACTCTTCCCGCTGTGGGT 368
353 CAAGAAGAGCAGGCGGAGTACATGAT 379

RESULT 6
AV392165 551 bp mRNA linear EST 23-APR-2002
LOCUS Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION cDNA clone CM083e05_r 5', mRNA sequence.
ACCESSION AV392165
VERSION AV392165.1 GI:6546381
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 551)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE A large scale structural analysis of cDNAs in a unicellular green
alga, Chlamydomonas reinhardtii. I. Generation of 3433
non-redundant expressed sequence tags
JOURNAL DNA Res. 6 (6), 369-373 (1999)
MEDLINE 20152988
PUBMED 10691129
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..551
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="C9"
/db_xref="taxon:3055"
/dev_stage="photoautotrophic growth"
/clone_lib="Chlamydomonas reinhardtii C9"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 94 a 182 c 189 g 85 t 1 others
ORIGIN
Query Match 5.6%; Score 40.6; DB 9; Length 551;
Best Local Similarity 45.3%; Pred. No. 8.2; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 179;

QY 42 CGACCTCGTGGGTACATTCGCTCGTGGCGCCCCCTAGGGGGCGTGCAGGGCCCT 101
108 CGAGCTCATCTCGTCAATTTGTGGCGGCACTGCCAACATGAAGGACGTGCTGACGACCT 167
102 GGGCGATGCGCTCCGGTTCTGAGGACGCGGTGAACATGCAACAGGAATTTCCCGG 161
168 GCGCGCGCGCGCGGAGTGGGAGGCGGCTACGCGCACGAGTCCGTGAGCTTGGGCGC 227
162 TTGCTCTTCTATCTCTTCTTGTCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 221
228 CCGCAAGGTGTTTGAAGATCAAGAGTACGTGTAACCTCAAGGCCAGAACCCAG 287
222 TTATGAAGTGCACAGCTGTCCGGATGTATACCATGTACGAACGATGCTCCCACTCAAG 281

```

```

288 CTTGCGCGTCCGCTGGTGGGCACTCGCTGGCGGCGGCACCGCGGCTGCTCTCGAT 347
282 CATTTGTATGAGGACGCGGATGATCATGTACACACCCCGGTCGTCGCTGCGTTTCG 341
348 CCTGATGCACACGACGAGGAGTTTGGCGGCGGATCTACGGCGGCGTGCCTATGCCGGG 407
342 GGAGAACAACTCTTCCCGCTGTGGGT 368
408 CAAGNAGAGCAAGGCGAGCTACATGAT 434

RESULT 7
BI996341 552 bp mRNA linear EST 25-OCT-2001
LOCUS 1031037A07.y2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI996341
VERSION BI996341.1 GI:16431115
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 552)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
JOURNAL Unpublished
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source
1..552
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hr); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
Research 6: 791-806."

BASE COUNT 93 a 184 c 189 g 86 t
ORIGIN
Query Match 5.6%; Score 40.6; DB 12; Length 552;
Best Local Similarity 45.3%; Pred. No. 8.2; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 179;

QY 42 CGACCTCGTGGGTACATTCGCTCGTGGCGCCCCCTAGGGGGCGTGCAGGGCCCT 101
110 CGAGCTCATCTCGTCAATTTGTGGCGGCACTGCCAACATGAAGGACGTGCTGACGACCT 169

```



```

QY 102 GGCGCATGGCTCCGGTTCAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCG 161
Db 170 GGCGCGCGCGCGCGAGTGGAGGCGCGCTACGGCACGAGTCCGTGAGCTTGGCGCG 229
QY 162 TTGCTCTTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 221
Db 230 CCGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCAAGACCCAG 289
QY 222 TTATGAAGTGGCAAGCTGTCGGGATGTACCATGTACGAAGAGCTGCTTCAACTCAAG 281
Db 290 CTTGCGCGCTCCGTGCTGGCGGCACCTCGCTGGCGCGGCACCGCGGCTGCGCTGCGAT 349
QY 282 CATTTGTTATGAGCAGCGACATGATCATGCACACCCCGGTGGTGGCTGCGTTCG 341
Db 350 CTGTATGACACAGAGAGTTTGGCGCGCATCTACGGCGGCTGCCCATGCCGGG 409
QY 342 GGAGAAACAATCTTCCCGCTGCTGGGT 368
Db 410 CAAGAAGAGCAAGGCGAGCTACATGAT 436

RESULT 8
BI727879
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1..584
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
Research 6: 791-806." 93 t
106 a 188 c 197 g

BASE COUNT
ORIGIN

Query Match 5.6%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. NO. 8.4;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTACATTCGCTCGTGGCGCCCGCTAGGGGGCGTCCAGGGCCCT 101
Db 47 CGAGCTCATCTCGTTCATTTGTGCGCGGCACTGCAACATGAAGAGCTGCTGACGACCT 106
QY 102 GGCGCATGGCTCCGGGTTCTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGG 161
Db 107 GGCGCGCGCGCGCGAGTGGGAGGGCGGTACGCGCAGAGTCCGTGAGCTTGGGCGC 166
QY 162 TTGCTCTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 221
Db 167 CCGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCAAGACCCAG 226
QY 222 TTATGAAGTGGCAACGCTGTCGGGATGTACCATGTACGAAGAGCTGCTTCAACTCAAG 281
Db 227 CTTGCGCGCTCCGTGCTGGCGGCACCTCGCTGGCGCGGCACCGCGGCTGCTGCGAT 286
QY 282 CATTTGTTATGAGCAGCGACATGATCATGCACACCCCGGTGGTGGCTGCGTTCG 341
Db 287 CTGTATGACACAGAGAGTTCGGCGCGCATCTACGGCGGCTGCCCATGCCGGG 346
QY 342 GGAGAAACAATCTTCCCGCTGCTGGGT 368
Db 347 CAAGAAGAGCAAGGCGAGCTACATGAT 373

RESULT 9
BX356664
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI015YB03"
/lisuse_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN

Query Match 5.6%; Score 40.4; DB 13; Length 1201;
Best Local Similarity 10.4%; Pred. No. 12;
Matches 52; Conservative 233; Mismatches 212; Indels 3; Gaps 1;

```

/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 224 a 499 c 240 g 198 t 1 others
ORIGIN

Query Match 5.6%; Score 40.2; DB 12; Length 1162;
Best Local Similarity 54.4%; Pred. No. 13;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 313 CACACCCCCGGGTGCGTCCCTCGTTCGGGAGACAACTTCCCGCTGGGTAGCG 372
Db 715 CCCCCCGGGTCCCTGCGCCACACCCCGCGACCCCAAAACCCCCCGGACCGCTCC 774
QY 373 CTCACCCCGAGCTCGAGCTAGGAAGCGGAGCGTCCACACGACAAATACGACGCCAC 432
Db 775 CCCACTGCCACCGACCCCGCCCATATCGCCCCCTACCGGATCACCCTACCCAC 834
QY 433 GTCGATTCGCCAGCTGTTCACCATCTCGCC 461
Db 835 GCCTGATCCCGCCCTGCACACCCCGCC 863

RESULT 11
AL513886 1201 bp mRNA linear EST 08-MAY-2003
LOCUS AL513886 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA006ZG08
DEFINITION 5-PRIME. mRNA sequence.
ACCESSION AL513886
VERSION AL513886.2 GI:30463771
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi.12777380.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4924.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BA006ZG08P1&cluster=4924.f. Contact :
Peng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BA006ZG08P1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA006ZG08"
/tissue_type="PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
201 a 311 c 349 g 146 t 194 others
BASE COUNT
ORIGIN

FEATURES
source

34 GCCTTCGGCACTCGTGGGTACATTCGCTGCGCGCCCGCTAGGGGCGCTGCC 93
Db 618 GNTNTSSSSSTNNNNSSSSSSNNNTNTTBTBTSSSTSSSTSSSTBTSSST 677
QY 94 AGGGCCCTGGCATCGGTCGGGTCTCGGAGGACGCGTGAATGACACAGGGAAT 153
Db 678 SSSSSBBTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 737
QY 154 TTGCCCGGTGCTCTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 213
Db 738 TTKSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 797
QY 214 GCTTCGCTTAAAGTGGCAACGTCGCGGAGTACCATGTCAGCAACGACGTGCC 273
Db 798 TBSMTSSSBTCTSSSSSSSBTSSSTSSSTSSSTSSSTSSSTSSSTSSST 854
QY 274 AACTCAAGCATTCGTATAGGCGAGCGACATGATCATCACACCCCGGTCGCGCC 333
Db 855 TSSTTNTTSSYSBSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 914
QY 334 TGGTTCGGGAGAACAACTTCCCGCTGCTGGGTAGCGCTCACCCACGCTCCGAGCT 393
Db 915 SSSSTTBSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 974
QY 394 AGNACCCAGCGTCCCGCACGACAAACGACGACGTCGATTCGCCAGCTGTTCACC 453
Db 975 STTBSSTTTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 1034
QY 454 ATCTCGCTCGCGCGATGAGCGTGCAGGATGCAATGTGCTCAATCTATCCCGGCCAC 513
Db 1035 SNNKSSSSSSSSSSSSSSSSSSSBTTTBTSSSTSSSTSSSTSSSTSSST 1094
QY 514 ATAACGGGTACCGTATGCC 533
Db 1095 SBTSSSTTTTSSSATBSB 1114

RESULT 10
BM918259 1162 bp mRNA linear EST 12-MAR-2002
LOCUS BM918259
DEFINITION AGENCOURT_6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649
5', mRNA sequence.
ACCESSION BM918259
VERSION BM918259.1 GI:19368638
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1162)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at:
http://image.llnl.gov
Plate: L1CM2016 row: n column: 18
High quality sequence stop: 567.
Location/Qualifiers
1. 1162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5485649"
/tissue_type="natural killer cells, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"

FEATURES
source

109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```
Query Match      5.5%; Score 40; DB 9; Length 1201;
Best Local Similarity 26.5%; Pred. No. 15;
Matches 103; Conservative 104; Mismatches 179; Indels 3; Gaps 1;

QY 32 GCGGCTTCGCCACCTCGTGGGTACATTCGCTCGCGCGCCCTAGGGGGGCTG 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1129 GSGCSGTGCSBSNCTYKKGKBSSSCCSVSSGSGSCSSCSGCCGGGGGGGG 1070
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 CAGGGCCCTCGCGCATGCGCTCGGGTTCGGAGACGCGGTGAACACTATGCAACAGGA 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1069 CCSCGKGGGGKSSSGSGGCCCGCGG--GGSSSGSGGGGSCCWAASSAYKKKK 1013
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 ATTGCGCGGTGCTCTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1012 GKKRTTTTMMAAATHTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 953
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 CAGCTTCGCTTATGAAGTCGCGCACTGTCGCGATGTACCATGTGTCAGCAAGACTGCT 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 952 GACSCSCGCAADAACVAGGMDGAMKSGTGVGSCCTTSRCKWGGGTSGMVGCATTYM 893
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 CCAACTCAAGCATTTGTATAGGACGCGACATGATCATGACACCCCGGGTGGCTGC 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 AYBSYTGTRRTWTGTTGCTGTATGAGSGMYSSKKBGKCCMAVACSCGAGASCST 833
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 CTGCGTTCGCGAGAACACTCTTCCGCTGCTGGTAGCGCTCACCCCGCTCGCAG 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 SCGSKGTTTCTTGTGCTGTGAAGSMBGRTWAGGGGGGGGGCCCYSCMCCCCCYB 773
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 CTAGGAACGCCAGCGTCCCGCCACCAACCA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 BBCCMCCHCTKCKWCCGACTYCCCCA 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
BJ252669/c
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
BJ252669
VERSION
BJ252669.1 GI:20061830
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 359)
Ogiwara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadaasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
BASE COUNT 70 a 115 c 107 g
ORIGIN
source

Query Match      5.4%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 18;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTCGCCGACCTCGTGGGTACATTCGCTCGCGCGCCCTAGG 83
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 CTTCAAGTCAACAGCGCGCTCTGGAAGCGGCTCAGGGCGGTGACGCCGCTCGCG 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 84 GGGCGCTGCGCAGGGCCCTGGCGCATGCGCTCGGGTTCGGAGGACGGGTGAACATG 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GGACGCCGCGCAGCCCTGGCGCAGGACGTCGACGCTGCCCGTGCACGTGCCCAAG 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
CC335916
LOCUS
DEFINITION
OGUJA160TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0393124,
genomic survey sequence.
ACCESSION
CC335916
VERSION
CC335916.1 GI:30805329
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Ogiwara, Y. and Murai, K.
Expressed genes in Zea mays
Unpublished
Contact: Tadaasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
BASE COUNT 70 a 115 c 107 g
ORIGIN
source
```

Mon Dec 22 13:28:39 2003

REFERENCE 1 (bases 1 to 840)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished
COMMENT Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

Location/Qualifiers

1. .840

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0393124"

/clone_lib="ZM 0.7 1.5 KB"

/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb"

methylation filtered genomic DNA library"

103 a 298 c 291 g 148 t

BASE COUNT

ORIGIN

Query Match 5.4%; Score 39; DB 29; Length 840;

Best Local Similarity 53.6%; Pred. No. 24;

Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 39 CGCGACCTCGTGGGTACATTCGGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGC 98

Db 292 CCCGGCGGTCCAGCGCGCCACTCCGCTTCACTGCGCGCTCCGGCGGCTCGCGGAACGC 351

Qy 99 CCTGGCGCATGGCGTCCGGGTTCCTGGAGGACGGCGTGAACATATGCAACAGGGGAATTGCC 158

Db 352 TGTGGCGGTCTCAAGCTGCTCTCTGGGTGCGCGCGGACCCGGTAACTCCCGACTCCGC 411

Qy 159 CGGTGGCTCTTCTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189

Db 412 CGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 442

Search completed: December 20, 2003, 06:54:48

Job time : 1862.96 secs

CD446071 873 bp mRNA linear EST 03-JUN-2003

LOCUS ELO1T0207B11.b Endosperm_4 Zea mays cDNA, mRNA sequence.

DEFINITION CD446071

ACCESSION CD446071.1 GI:31361714

VERSION EST.

KEYWORDS Zea mays

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 873)

Lai,J., Dev,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and

Messing,J.

Sequencing of the maize endosperm ESTs

Unpublished

Contact: Lai, Jinsheng

Dr. Joachim Messing's lab

Waksman Institute, Rutgers University

190 Frelinghuysen Rd., Piscataway, NJ 08854, USA

Tel: 732-445-3801

Fax: 732-445-5735

Email: jlai@waksman.rutgers.edu

Seq primer: T3.

Location/Qualifiers

1. .873

/organism="Zea mays"

/mol_type="mRNA"

FEATURES

source

109 a 300 c 294 g 170 t

BASE COUNT

ORIGIN

Query Match 5.4%; Score 39; DB 14; Length 873;

Best Local Similarity 53.6%; Pred. No. 24;

Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 39 CGCGACCTCGTGGGTACATTCGGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGC 98

Db 583 CCCGGCGGTCCAGCGCGCCACTCCGCTTCACTGCGCGCTCCGGCGGCTCGCGGAACGC 642

Qy 99 CCTGGCGCATGGCGTCCGGGTTCCTGGAGGACGGCGTGAACATATGCAACAGGGGAATTGCC 158

Db 643 TGTGGCGGTCTCAAGCTGCTCTCTGGGTGCGCGCGGACCCGGTAACTCCCGACTCCGC 702

Qy 159 CGGTGGCTCTTCTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189

Db 703 CGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 733

/cultivar="W22"

/db_xref="taxon:4577"

/tissue_type="Endosperm of 7-23DAP"

/clone_lib="Endosperm_4"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI"

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 52.1425 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-21
Perfect score: 723
Sequence: 1 ATGTTGGTAGGTGATCGA.....TACTCTTTGCTCCCTAATAG 723

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723	100.0	723	3	US-08-612-973-21
2	723	100.0	723	3	US-08-927-597-21
3	641	88.7	795	3	US-08-612-973-5
4	641	88.7	795	3	US-08-927-597-5
5	624.8	86.4	2082	3	US-08-612-973-47
6	624.8	86.4	2082	3	US-08-927-597-47
7	624.8	86.4	2433	3	US-08-612-973-49
8	624.8	86.4	2433	3	US-08-927-597-49
9	598.2	82.7	606	3	US-08-612-973-25
10	598.2	82.7	606	3	US-08-927-597-25
11	597	82.6	636	3	US-08-612-973-27
12	597	82.6	636	3	US-08-927-597-27
13	556	76.9	561	3	US-08-612-973-23
14	556	76.9	561	3	US-08-927-597-23
15	549.6	76.0	1539	2	US-08-470-426B-17
16	549.6	76.0	1863	2	US-08-470-426B-14
17	549.6	76.0	2116	3	US-08-191-160-21
18	548	75.8	932	1	US-08-081-072-15
19	548	75.8	932	1	US-08-449-093A-15
20	548	75.8	9395	3	US-09-014-416-4
21	548	75.8	9599	3	US-09-014-416-6
22	540	74.7	9472	4	US-08-150-204E-96
23	534.6	73.9	1167	2	US-08-324-977-9
24	534.6	73.9	1167	2	US-08-384-616-9
25	534.6	73.9	1167	2	US-08-904-686A-9
26	534.6	73.9	1167	3	US-09-315-850-9
27	534.6	73.9	1499	1	US-08-324-977-3

28	534.6	73.9	1499	2	US-08-384-616-3	Sequence 3, Appl
29	534.6	73.9	1499	2	US-08-904-686A-3	Sequence 3, Appl
30	534.6	73.9	1499	3	US-09-315-850-3	Sequence 3, Appl
31	534.6	73.9	6039	1	US-08-324-977-11	Sequence 11, Appl
32	534.6	73.9	6039	2	US-08-384-616-11	Sequence 11, Appl
33	534.6	73.9	6039	3	US-08-904-686A-11	Sequence 11, Appl
34	534.6	73.9	6039	3	US-09-315-850-11	Sequence 11, Appl
35	534.6	73.9	9030	1	US-08-324-977-13	Sequence 13, Appl
36	534.6	73.9	9030	2	US-08-384-616-13	Sequence 13, Appl
37	534.6	73.9	9030	2	US-08-904-686A-13	Sequence 13, Appl
38	534.6	73.9	9030	3	US-09-315-850-13	Sequence 13, Appl
39	534.6	73.9	9416	1	US-08-324-977-1	Sequence 1, Appl
40	534.6	73.9	9416	2	US-08-384-616-1	Sequence 1, Appl
41	534.6	73.9	9416	2	US-08-904-686A-1	Sequence 1, Appl
42	534.6	73.9	9416	3	US-09-315-850-1	Sequence 1, Appl
43	534.6	73.9	9416	4	US-08-823-895A-27	Sequence 27, Appl
44	493.2	68.2	742	1	US-08-081-072-18	Sequence 18, Appl
45	493.2	68.2	742	1	US-08-449-093A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-21
; Sequence 21, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: 1..717

US-08-612-973-21

	Query Match	100.0%;	Score 723;	DB 3;	Length 723;
	Best Local Similarity	100.0%;	Pred. No. 2.5e-183;		
	Matches 723;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGTTGGGTAGTCAATCGATACCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT	60		
DB	1	ATGTTGGGTAGTCAATCGATACCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT	60		
QY	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120		
DB	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120		
QY	121	CTGGAGGACGGCGTGAATCTATGCAACAGGGAATTCGCCGGTTGCTTTCTATCTTC	180		
DB	121	CTGGAGGACGGCGTGAATCTATGCAACAGGGAATTCGCCGGTTGCTTTCTATCTTC	180		
QY	181	CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTTAAGTGGCGAACGTTG	240		
DB	181	CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTTAAGTGGCGAACGTTG	240		
QY	241	TCGGGATGTACATGTCTAGCAACGACGTGCCAACTCAAGCAATGTGTATCAGGCAGCG	300		
DB	241	TCGGGATGTACATGTCTAGCAACGACGTGCCAACTCAAGCAATGTGTATCAGGCAGCG	300		
QY	301	GACATGATCATGCACACCCCGGGTGGCTGCCCTCGCTTCGGGAGAACAACTCTTCCCGC	360		
DB	301	GACATGATCATGCACACCCCGGGTGGCTGCCCTCGCTTCGGGAGAACAACTCTTCCCGC	360		
QY	361	TGCTGGGTAGCGTCAACCCGACGCTCGCAGTATAGGACGCCAGCGTCCCGACACGACA	420		
DB	361	TGCTGGGTAGCGTCAACCCGACGCTCGCAGTATAGGACGCCAGCGTCCCGACACGACA	420		
QY	421	ATACGAGCCACGTCGATTCACCATTCGCTCGCGGATGAGACGGTGT	480		
DB	421	ATACGAGCCACGTCGATTCACCATTCGCTCGCGGATGAGACGGTGT	480		
QY	481	CAGGATGCAATTTGCTCAATCTATCCGGCGACATAACGGGTACCGGTATGGGAT	540		
DB	481	CAGGATGCAATTTGCTCAATCTATCCGGCGACATAACGGGTATGGGAT	540		
QY	541	ATGATGATGAATCTGGTCCCTACAAAGGCCCTGGTGTATCGAGCTGCTCCGATCCCA	600		
DB	541	ATGATGATGAATCTGGTCCCTACAAAGGCCCTGGTGTATCGAGCTGCTCCGATCCCA	600		
QY	601	CAAGCTGCTGGGACATGGTGGGGGCCCATTTGGGAGTCTCGGGGTCTCGCCTAC	660		
DB	601	CAAGCTGCTGGGACATGGTGGGGGCCCATTTGGGAGTCTCGGGGTCTCGCCTAC	660		
QY	661	TATTTCCATGGTGGGAACTGGGCTTAAGTTTTGATTTGTGATGCTACTCTTTGCTCCCTAA	720		
DB	661	TATTTCCATGGTGGGAACTGGGCTTAAGTTTTGATTTGTGATGCTACTCTTTGCTCCCTAA	720		
QY	721	TAG 723			
DB	721	TAG 723			

RESULT 2
US-08-527-597-21
; Sequence 21, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEBERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: EUSEE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.

STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..720
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..717
US-08-927-597-21

Query Match	100.0%;	Score 723;	DB 3;	Length 723;
Best Local Similarity	100.0%;	Pred. No. 2.5e-183;		
Matches 723;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGTTGGGTAAAGTCATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGGTACATT	60	
Db	1	ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGGTACATT	60	
Qy	61	CCGCTCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCGGGGT	120	
Db	61	CCGCTCTGTGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCGGGGT	120	
Qy	121	CTGAGAGCGCGGTGAACATGCAACAGGGAAATTTGCCGGTTCGTCTTTCTATCTTC	180	
Db	121	CTGAGGAGCGCGGTGAACATGCAACAGGGAAATTTGCCGGTTCGTCTTTCTATCTTC	180	
Qy	181	CTCTTGGCTTTTGCTGTCTCTGTCTGACCGTTTCAGCTTCGGTTCGATGAGTGGCAACGTG	240	
Db	181	CTCTTGGCTTTTGCTGTCTCTGTCTGACCGTTTCAGCTTCGGTTCGATGAGTGGCAACGTG	240	
Qy	241	TCCGGGATGTACCATGTTCAGCAACGACGTCTCCAACCTCAAGCATTTGTGTATGAGCAGCG	300	
Db	241	TCCGGGATGTACCATGTTCAGCAACGACGTCTCCAACCTCAAGCATTTGTGTATGAGCAGCG	300	
Qy	301	GACATGATCATGCACACCCCGGGTGCGTGCCCTCGGTTCCGGAGAACAACTCTTCCCGC	360	
Db	301	GACATGATCATGCACACCCCGGGTGCGTGCCCTCGGTTCCGGAGAACAACTCTTCCCGC	360	
Qy	361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTTAGGAAAGCCAGCGTCCCCACCACGACA	420	
Db	361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTTAGGAAAGCCAGCGTCCCCACCACGACA	420	

QY 421 ATACGACGCGCAGCTCGATTCCAGCTGTTCAACCATCTCGCTCGCGGATGAGCGTG 480
DB 421 ATACGACGCGCAGCTCGATTCCAGCTGTTCAACCATCTCGCTCGCGGATGAGCGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGGTATGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGGTATGGGAT 540
QY 541 ATGATGATGAAGTGGTGGCTCAACAGCGCCCTGGTGTATCGAGCTGCTCCGGATCCCA 600
DB 541 ATGATGATGAAGTGGTGGCTCAACAGCGCCCTGGTGTATCGAGCTGCTCCGGATCCCA 600
QY 601 CAACTGCTGCTGACATGTTGGGGGGCCCAATTTGGGGAGTCTTGGGGTCTCGCTAC 660
DB 601 CAACTGCTGCTGACATGTTGGGGGGCCCAATTTGGGGAGTCTTGGGGTCTCGCTAC 660
QY 661 TATTCATGTTGGGGAACCTGGGCTAAGGTTTTCATTGTGATGCTACTCTTTGCTCCCTAA 720
DB 661 TATTCATGTTGGGGAACCTGGGCTAAGGTTTTCATTGTGATGCTACTCTTTGCTCCCTAA 720
QY 721 TAG 723
DB 721 TAG 723

RESULT 3

US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 1..789
US-08-612-973-5
Query Match 88.7%; Score 641; DB 3; Length 795;
Best Local Similarity 90.9%; Pred. No. 1.7e-161;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGGTCAATCGATACCTTATCATGCGGCTTCGCCAACCCTCGTGGGATCAATT 60
DB 1 ATGTTGGGTAAGGTCAATCGATACCTTATCATGCGGCTTCGCCAACCCTCGTGGGATCAATT 60
QY 61 CGCTCGTGGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGGCGATGCGGCTCCGGGTT 120
DB 61 CGCTCGTGGGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGGCGATGCGGCTCCGGGTT 120
QY 121 CTGGAGAGCGGCGTGAACATATGCAACAGGGAAATTTGCCCGGTTTCTTTCTCTATCTTC 180
DB 121 CTGGAGAGCGGCGTGAACATATGCAACAGGGAAATTTGCCCGGTTTCTTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGTCTGACCGTTTCCAGCTTTCGCTTATGAAGTGCACAGTG 240
DB 181 CTCTTGGCTTTGCTGCTGCTGTCTGACCGTTTCCAGCTTTCGCTTATGAAGTGCACAGTG 240
QY 241 TCCGGGATGTACCATGTCTACGAAACGATGCTTCCAACTCAAGCAATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTCTACGAAACGATGCTTCCAACTCAAGCAATTTGTATGAGGACGG 300
QY 301 GACATGATCATGACACACCCCGGTCGCTGCGCTTCCAGCTTTCGCTTATGAAGTGCACAGTG 360
DB 301 GACATGATCATGACACACCCCGGTCGCTGCGCTTCCAGCTTTCGCTTATGAAGTGCACAGTG 360
QY 361 TGCTGGGTAGCGCTCACCCCGGTCGCTGCGCTTCCAGCTTTCGCTTATGAAGTGCACAGTG 420
DB 361 TGCTGGGTAGCGCTCACCCCGGTCGCTGCGCTTCCAGCTTTCGCTTATGAAGTGCACAGTG 420
QY 421 ATAGAGCGGACGTCGAT
DB 421 ATAGAGCGGACGTCGAT
QY 439 -----TCCCAGCTGTTCAACCATCTCGCTCCCGG 468
DB 481 GGGGACCTCTGCGGATCTGTTCTCTGCTTCCAGCTGTTCAACCATCTCGCTCCCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAAGCGGTCAACGT 528
DB 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATAAGCGGTCAACGT 600
QY 529 ATGGCTTGGGATATGATGATGAACCTGGTCCCTTACAAACGCGCCCTGGTGTATCGCAGCTG 588
DB 601 ATGGCTTGGGATATGATGATGAACCTGGTCCCTTACAAACGCGCCCTGGTGTATCGCAGCTG 660
QY 589 CTCGGGATCCCAAGCTGTCGTGGACATGTTGGGGGGGCGCCCATTTGGGGAGTCTCTGGCG 648
DB 661 CTCGGGATCCCAAGCTGTCGTGGACATGTTGGGGGGGCGCCCATTTGGGGAGTCTCTGGCG 720
QY 649 GGTCTCGCTTACTATTTCATGTTGGGGAACCTGGGCTTAAGGTTTGTATGTTGATGCTACTC 708
DB 721 GGTCTCGCTTACTATTTCATGTTGGGGAACCTGGGCTTAAGGTTTGTATGTTGATGCTACTC 780
QY 709 TTTGCTCCCTAATAG 723
DB 781 TTTGCTCCCTAATAG 795
RESULT 4
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY

Mon Dec 22 13:28:37 2003

APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF INVENTION: 111
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789

US-08-927-597-5
Query Match 88.7%; Score 641; DB 3; Length 795;
Best Local Similarity 90.9%; Pred. No. 1.7e-161;
Matches 723; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCCCTTACGAGGCGCTGCAGGCGCTTCGCCGACCTCGTGGGT 120
DB 61 CCGCTCGTGGGCCCTTACGAGGCGCTGCAGGCGCTTCGCCGACCTCGTGGGT 120
QY 121 CTGGAGGACGGCTGAATATGCAACAGGGAATTTGCCCGGTGCTTTCTATCTTC 180
DB 121 CTGGAGGACGGCTGAATATGCAACAGGGAATTTGCCCGGTGCTTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGACGGTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 240
DB 181 CTCTTGGCTTTGCTGCTCTGACGGTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTC 240
QY 241 TCCGGGATGACCATGTCACGACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGACCATGTCACGACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTTCCGGGAGCAACTCTTCCCGC 360

Db 301 GACATGATCATGCACACCCCGGGTGGCTTCCGGGAGCAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCAACCGCTCGAGCTAGGAACGCGCAGCGTCCCACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCAACCGCTCGAGCTAGGAACGCGCAGCGTCCCACACGACA 420
QY 421 ATACGACGCCAGCTGAT-----TCCAGCTGTTCCAGCATCTCGCTCGCGG 438
Db 421 ATACGACGCCAGCTGATTTGCTGTTGGGGGGGCTGCTTTCTGTTCCGCTAATGACGTG 480
QY 439 -----TCCAGCTGTTCCAGCATCTCGCTCGCGG 468
Db 481 GGGGACCTCTCGGATCTGCTTCTCGCTTCCAGCTGTTCCAGCATCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTCCTATCCCGGCCACATTAACGGGTACCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTCCTATCCCGGCCACATTAACGGGTACCGT 600
QY 529 ATGGCTTGGGATATGATGATGAATGCTGCTGCTCAACAGCGCCCTGCTGATCGCAGCTG 588
Db 601 ATGGCTTGGGATATGATGATGAATGCTGCTGCTCAACAGCGCCCTGCTGATCGCAGCTG 660
QY 589 CTCGGATCCCAAGCTGCTGAGCATGCTGGGGGGGCGCCATTTGGGGAGTCTCTGGCG 648
Db 661 CTCGGATCCCAAGCTGCTGAGCATGCTGGGGGGGCGCCATTTGGGGAGTCTCTGGCG 720
QY 649 GGTCTCGCTACTATTCCATGCTGGGGAACCTGGGCTAAAGTTTGTATGATGCTACTC 708
Db 721 GGTCTCGCTACTATTCCATGCTGGGGAACCTGGGCTAAAGTTTGTATGATGCTACTC 780
QY 709 TTTGCTCCCTAATAG 723
Db 781 TTTGCTCCCTAATAG 795

RESULT 5
US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTINOFF, GUY
; APPLICANT: BUYS, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs

Db 1015 CGGATCCCAAGCTGTCTGGGACATGGTGGCGGGGCCCAATTGGGGAGTCTTGGCGGGC 1074
QY 652 CTGCGCTACTATTCCATGTGGGGAACCTGGGCTTAAGGTTTGTGATGTGATCTCTTT 711
Db 1075 CTGCGCTACTATTCCATGTGGGGAACCTGGGCTTAAGGTTTGTGATGTGATCTCTTT 1134
QY 712 GC 713
Db 1135 GC 1136

RESULT 8

US-08-927-597-49
; Sequence 49, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..2427
US-08-927-597-49

Query Match 86.4%; Score 624.8; DB 3; Length 2433;
Best Local Similarity 90.5%; Pred. No. 4.9e-157;
Matches 708; Conservative 2; Indels 72; Gaps 1;
QY 4 TTGGGTAAGTCAATCCCTTACATCGCGGCTTCGCGGACCTCGTGGGTACATCCG 63
|||||

Db 355 TTGGGTAAGTCAATCCCTTACATCGCGGCTTCGCGGACCTCGTGGGTACATCCG 414
QY 64 CTCGTGGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTCTTG 123
Db 415 CTCGTGGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTCTTG 474
QY 124 GAGGACGGGTGAACACTATGCAACAGGGAATTTGCGCGGTGCTCTTTCTCTATCTTCCTC 183
Db 475 GAGGACGGGTGAACACTATGCAACAGGGAATTTGCGCGGTGCTCTTTCTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTCGCGCAACGTGTC 243
Db 535 TTGGCTTTGCTGCTGCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTCGCGCAACGTGTC 594
QY 244 GGGATGTACCATGTGTACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCGAGGGAC 303
Db 595 GGGATGTACCATGTGTACGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGCGAGGGAC 654
QY 304 ATGATCATGCACACCCCCGGGTGCGTCCCTGCGTTCGGGAGAAACAACTCTTCCCGCTGC 363
Db 655 ATGATCATGCACACCCCCGGGTGCGTCCCTGCGTTCGGGAGAAACAACTCTTCCCGCTGC 714
QY 364 TGGGTAGCGCTCACCCCCGACTCGCAGCTAGGAACGCCAGCGTCCCAACCCACGACAATA 423
Db 715 TGGGTAGCGCTCACCCCCGACTCGCAGCTAGGAACGCCAGCGTCCCAACCCACGACAATA 774
QY 424 CGAGCCGACTCGAT----- 438
Db 775 CGAGCCGACTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGG 834
QY 439 -----TCCAGCTGTTCAACATCTCGCTCGCGGCGCAT 471
Db 835 GACCTCTGCGGATCTGCTTCTCTCCAGCTGTTCAACATCTCGCTCGCGGCGCAT 894
QY 472 GAGCGGTGCGGACTGCAATTCCTCAATCTATCCGCGCACATAACGGGTACCGCTATG 531
Db 895 GAGCGGTGCGGACTGCAATTCCTCAATCTATCCGCGCACATAACGGGTACCGCTATG 954
QY 532 GCTTGGGATATGATGATGAACCTGCTCGCTCAAAACGGCCCTGGTGTATCGAGCTGCTC 591
Db 955 GCTTGGGATATGATGATGAACCTGCTCGCTCAAAACGGCCCTGGTGTATCGAGCTGCTC 1014
QY 592 CGGATCCCAAGCTGCTGTGACATGTGTGGGGGGGCCCATTTGGGGAGTCTTGGCGGGT 651
Db 1015 CGGATCCCAAGCTGCTGTGACATGTGTGGGGGGGCCCATTTGGGGAGTCTTGGCGGGC 1074
QY 652 CTGCGCTACTATTCCATGTGGGGAACCTGGGCTTAAGGTTTGTGATGTGATCTCTTT 711
Db 1075 CTGCGCTACTATTCCATGTGGGGAACCTGGGCTTAAGGTTTGTGATGTGATCTCTTT 1134
QY 712 GC 713
Db 1135 GC 1136

RESULT 9

US-08-612-973-25
; Sequence 25, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.

us-09-899-303a-21.rni

Mon Dec 22 13:28:37 2003

541 ATGATGATCAACTGGTCGGCTTCAACAGCGCCCTGGTGGTATCGAGCTGCTCGGATCCCA 600
541 ATGATGATCAACTGGTCGGCTTCAACAGCGCCCTGGTGGTATCGAGCTGCTCGGATCCCTC 600
601 CAA 603
601 TAA 603

RESULT 10
US-08-927-597-25
; Sequence 25, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..603
; NAME/KEY: mat peptide
; LOCATION: 1..600
; US-08-927-597-25

Query Match 82.7%; Score 598.2; DB 3; Length 606;
Best Local Similarity 99.5%; Pred. No. 3.8e-150;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTGGGTAAAGTCACTGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTGGGTAAAGTCACTGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..603
NAME/KEY: mat peptide
LOCATION: 1..600
US-08-612-973-25
Query Match 82.7%; Score 598.2; DB 3; Length 606;
Best Local Similarity 99.5%; Pred. No. 3.8e-150;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTGGGTAAAGTCACTGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTGGGTAAAGTCACTGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTCGCGGCCCTTAGGGGGCGCTCCAGAGGCCCTGGCGCATGCGGTCGGGTT 120
DB 61 CCGCTCGTCGCGGCCCTTAGGGGGCGCTCCAGAGGCCCTGGCGCATGCGGTCGGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCCTTGGCTTTGCTGCTGCTGTGACCGTTCCAGTTCGGTTTATGAAGTGGCAAGTG 240
DB 181 CTCCTTGGCTTTGCTGCTGCTGTGACCGTTCCAGTTCGGTTTATGAAGTGGCAAGTG 240
QY 241 TCCGGATGATACCATGTCAAGAACACTGCTCCAACTCAAGCATTTGTTATGAGCGAGCG 300
DB 241 TCCGGATGATACCATGTCAAGAACACTGCTCCAACTCAAGCATTTGTTATGAGCGAGCG 300
QY 301 GACATGATCATGACACACCCCGGGTTCGGTCCCTCGTTCGGGAGAACTCTTCCGCG 360
DB 301 GACATGATCATGACACACCCCGGGTTCGGTCCCTCGTTCGGGAGAACTCTTCCGCG 360
QY 361 TGTGGGTAGCGTCAACCCCGAGCTCCGAGTGAAGAGCGGAGTCCCAACACGACA 420
DB 361 TGTGGGTAGCGTCAACCCCGAGCTCCGAGTGAAGAGCGGAGTCCCAACACGACA 420
QY 421 ATACGACGCGCAGCTCGATTCCAGCTGTTTACCATTCTCGCTTCGGCGGATGAGCGGTG 480
DB 421 ATACGACGCGCAGCTCGATTCCAGCTGTTTACCATTCTCGCTTCGGCGGATGAGCGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGGCAATTAACGGGTCCAGCTATGCTTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCGGGCAATTAACGGGTCCAGCTATGCTTTGGGAT 540

QY 61 CCGCTGCTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGGT 120
Db 61 CCGCTGCTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGGT 120
QY 121 CTGGAGGAGCGGCTGAACATATGCAACAGGGAATTTGCCGGGTTGCTCTTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACATATGCAACAGGGAATTTGCCGGGTTGCTCTTTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGCTGCTGTCTGACCGGTTCCAGCTTCCGCTTATGAAGTGCACACG 240
Db 181 CTCCTGGCTTTGCTGCTGCTGTCTGACCGGTTCCAGCTTCCGCTTATGAAGTGCACACG 240
QY 241 TCCGGGATGATCATGCTGACGAGACGACTGCTCCNACTCAAGCATTTGTATGAGGACG 300
Db 241 TCCGGGATGATCATGCTGACGAGACGACTGCTCCNACTCAAGCATTTGTATGAGGACG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GACATGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Db 361 TCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
QY 421 ATACGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 ATACGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 CAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 ATGATGATGAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
Db 541 ATGATGATGAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
QY 601 CAA 603
Db 601 TAA 603

RESULT 11
US-08-612-973-27
; Sequence 27, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630
US-08-612-973-27

Query Match 82.6%; Score 597; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.2e-150;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGGTATCATGATACCTTATCATGGGGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGGTATCATGATACCTTATCATGGGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTGCTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGGT 120
Db 61 CCGCTGCTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTTCGGGGT 120
QY 121 CTGGAGGAGCGGCTGAACTATGCAACAGGGAATTTGCCGGGTTGCTCTTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACTATGCAACAGGGAATTTGCCGGGTTGCTCTTTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 CTCCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGGATGATCATGCTGACGAGACGACTGCTCCNACTCAAGCATTTGTATGAGGACG 300
Db 241 TCCGGGATGATCATGCTGACGAGACGACTGCTCCNACTCAAGCATTTGTATGAGGACG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GACATGATCATGACACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Db 361 TCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
QY 421 ATACGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 ATACGAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 CAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 ATGATGATGAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 597
Db 541 ATGATGATGAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 597

RESULT 12
US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY

301	GACATGATCATGCACACCCCCCGGGTGCCTCGCTTGGGAGAACAACTCTTCCGCG	360
	Db	
361	TGCTGGGTAGCGTCAACCCCGCCAGCTCGCAGCTAGGAACGCCAGCGTCCCCACCAACGACA	420
	Qy	
361		
361	TGCTGGGTAGCGTCAACCCCGCCAGCTCGCAGCTAGGAACGCCAGCGTCCCCACCAACGACA	420
	Db	
421	ATACGACGCCACGTCGATNTCCAGCTGTTTCCATCTCGCTCGCGGCATGAGACGGTG	480
	Qy	
421		
421	ATACGACGCCACGTCGATNTCCAGCTGTTTCCATCTCGCTCGCGGCATGAGACGGTG	480
	Db	
481	CAGAGCTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGCTTGGGAT	540
	Qy	
481		
481	CAGAGCTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGCTTGGGAT	540
	Db	
541	ATGATGATGAACCTGGTCGCCTACACGGCCCTGGTGATCGCAGCTGCTCCGGATC	597
	Qy	
541		
541	ATGATGATGAACCTGGTCGCCTACACGGCCCTGGTGATCGCAGCTGCTCCGGATC	597
	Db	

```

RESULT 13
US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; FOR DIAGNOSTIC AND THERAPEUTIC USE

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0.
Version #1.25 (BPO)

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER:  US/08/612,973
FILING DATE:  11-MAR-1996
CLASSIFICATION:  435
ATTORNEY/AGENT INFORMATION:
NAME:  BYRNE, THOMAS E
REGISTRATION NUMBER:  32,205
REFERENCE/DOCKET NUMBER:  1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE:  (703) 816-4000
TELEFAX:  (703) 816-4100
INFORMATION FOR SEQ ID NO:  23:
SEQUENCE CHARACTERISTICS:
LENGTH:  561 base pairs
TYPE:  nucleic acid
STRANDEDNESS:  single
TOPOLOGY:  linear
MOLECULE TYPE:  cdna
HYPOTHETICAL:  NO
ANTI-SENSE:  NO
FEATURE:
NAME/KEY:  CDS
LOCATION:  1..558
FEATURE:
NAME/KEY:  mat_peptide
LOCATION:  1..555
US-08-612-973-23

```

Query Match	76.9%;	Score 556;	DB 3;	Length 561;
Best Local Similarity	100.0%;	Pred. No. 6.4e-139;		

Query Match	76.0%;	Score 549.5;	DB 2;	Length 1539;
Best Local Similarity	84.5%;	Prod. No. 4.4e-137;		
Matches 661;	Conservative	0;	Mismatches 49;	Indels 72; Gaps 1;
QY	4	TTGGGTAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCGG	63	
DB	355	TTGGGTAAGGTCATCGATACCTTACATCGCGCTTCGCCGATCTCATGGGTATATTCCC	414	
QY	64	CTCGTCGGCGCCCCCTTAGGGGGCGCTCCAGGGGCCCTGGGGCGATGGCGTTCGGGTTCTG	123	
DB	415	CTCGTCGGCGCCCCCTTAGGGGGCGCTGCAGGGGCTTTGGCACACG3GTTCGGGTTCTG	474	
QY	124	GAGGACGGGTCGAATATGCAACAGG3AATTGCCCGGTTCGCTCTTTCTCTATCTTCCTC	183	
DB	475	GAGGACGGGCGTGAACATATGCAACAGG3AACTTGGCCGGTTGCTCTTTCTCTATCTTCCTC	534	
QY	184	TTGGCTTTTGCTGTCTGTCTGCTGACCGCTTCAGCGTTTCGGCTTTATGAAGTGGCAACGTGTCC	243	
DB	535	TTGGCTTTTGCTGTCTGTGTTTGACCATCCACGCTTCGGCTTATGAAGTGGCAACGTGTCC	594	
QY	244	GGGATGTACCATGTGTACGAAAGGACTGCTCCAACTCAAGCATTTGTGTATGAGCGACGGAC	303	
DB	595	GGGATATACCATGTGTACGAAAGGACTGCTCCAACTCAAGCATTTGTGTATGAGCGACGGAC	654	
QY	304	ATGATCATCGACACCCCGGGTCGTGCCCTCGGTTTCGGGAGAAACAACTCTTCCCGCTGC	363	
DB	655	ATGATCATGCTACTCTCCGGGTGCTGCCCTCGGTTTCGGGAGAACAAAGCTCCCGTTGC	714	
QY	364	TGGGTAGCGCTCACCCCCACGCTTCGCAGCTTAGGAAAGCCACGCGTCCCGACCGACAATA	423	

Search completed: December 20, 2003, 07:03:07
Job time : 56.1425 secs

Search completed: December 20, 2003, 07:03:07
Job time : 56.1425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2124.1 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-23
Perfect score: 561
Sequence: 1 ATGTTGGGTAAAGTCATCGA.....TCATGTCAGCTGCTAATAG 561

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb.ba.*	GenEmbl.*
2:	gb.hcg.*	
3:	gb.in.*	
4:	gb.om.*	
5:	gb.ov.*	
6:	gb.pat.*	
7:	gb.ph.*	
8:	gb.pl.*	
9:	gb.pr.*	
10:	gb.ro.*	
11:	gb.sts.*	
12:	gb.sv.*	
13:	gb.un.*	
14:	gb.vi.*	
15:	em.ba.*	
16:	em.fun.*	
17:	em.hum.*	
18:	em.in.*	
19:	em.mu.*	
20:	em.om.*	
21:	em.or.*	
22:	em.ov.*	
23:	em.pat.*	
24:	em.ph.*	
25:	em.pl.*	
26:	em.ro.*	
27:	em.sts.*	
28:	em.un.*	
29:	em.vi.*	
30:	em.hcg.hum.*	
31:	em.hcg.inv.*	
32:	em.hcg.other.*	
33:	em.hcg.mus.*	
34:	em.hcg.pln.*	
35:	em.hcg.rod.*	
36:	em.hcg.mam.*	
37:	em.hcg.vrt.*	
38:	em.sv.*	
39:	em.hcgo.hum.*	
40:	em.hcgo.mus.*	
41:	em.hcgo.other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	561	100.0	561	6	A48685	A48685 Sequence 23
2	561	100.0	561	6	AR157337	AR157337 Sequence
3	561	100.0	561	6	AX452772	AX452772 Sequence
4	561	100.0	561	6	AX685024	AX685024 Sequence
5	556	99.1	606	6	A48687	A48687 Sequence 25
6	556	99.1	606	6	AR157338	AR157338 Sequence
7	556	99.1	606	6	AX452774	AX452774 Sequence
8	556	99.1	606	6	AX685026	AX685026 Sequence
9	556	99.1	636	6	A48689	A48689 Sequence 27
10	556	99.1	636	6	AR157339	AR157339 Sequence
11	556	99.1	636	6	AX452776	AX452776 Sequence
12	556	99.1	636	6	AX685028	AX685028 Sequence
13	556	99.1	723	6	A48683	A48683 Sequence 21
14	556	99.1	723	6	AR157336	AR157336 Sequence
15	556	99.1	723	6	AX452770	AX452770 Sequence
16	556	99.1	723	6	AX685022	AX685022 Sequence
17	474	84.5	795	6	A48667	A48667 Sequence 5
18	474	84.5	795	6	AR157325	AR157325 Sequence
19	474	84.5	795	6	AX452754	AX452754 Sequence
20	474	84.5	795	6	AX685006	AX685006 Sequence
21	471	84.0	2082	6	A48709	A48709 Sequence 47
22	471	84.0	2082	6	AR157350	AR157350 Sequence
23	471	84.0	2082	6	AX452796	AX452796 Sequence
24	471	84.0	2082	6	AX685048	AX685048 Sequence
25	471	84.0	2433	6	A48711	A48711 Sequence 49
26	471	84.0	2433	6	AR157351	AR157351 Sequence
27	471	84.0	2433	6	AX452798	AX452798 Sequence
28	471	84.0	2433	6	AX685050	AX685050 Sequence 7
29	461.4	82.2	633	6	A48669	A48669 Sequence
30	461.4	82.2	633	6	AR157326	AR157326 Sequence
31	461.4	82.2	633	6	AX452756	AX452756 Sequence
32	461.4	82.2	633	6	AX685008	AX685008 Sequence
33	454	80.9	636	6	A48675	A48675 Sequence 13
34	454	80.9	636	6	AR157329	AR157329 Sequence
35	454	80.9	636	6	AX452762	AX452762 Sequence
36	454	80.9	636	6	AX685014	AX685014 Sequence
37	426.6	76.0	9379	14	AF165052	AF165052 Hepatitis
38	426.6	76.0	9379	14	AF207766	AF207766 Hepatitis
39	426.6	76.0	9379	14	HCVPOLYP	AJ000009 Hepatitis
40	423.4	75.5	3296	14	AB008441	AB008441 Hepatitis
41	423.4	75.5	3296	14	AB008442	AB008442 Hepatitis
42	423.4	75.5	3296	14	AB008443	AB008443 Hepatitis
43	423.4	75.5	9344	14	AB049096	AB049096 Hepatitis
44	423.4	75.5	9379	14	AF165051	AF165051 Hepatitis
45	421.8	75.2	9359	14	AF313916	AF313916 Hepatitis

ALIGNMENTS

RESULT 1
A48685
LOCUS A48685 Sequence 23 from Patent WO9604385. 561 bp DNA linear PAT 07-MAR-1997
DEFINITION A48685
ACCESSION A48685
VERSION A48685.1 GI:2302398
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 561)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buysse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 23 15-FEB-1996;

Mon Dec 22 13:28:39 2003

INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
source
Location/Qualifiers
1..561
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
CDS
1..558
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03135.1"
/db_xref="GI:2302399"
VNYATGNLPGCCSFSLFALLSLCLTPASAEVNVNMGVHVTNDSCSSVYRAADM
IMHTPCGVCPCVRNENSSRCWALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSYFPHGTGHRMADMMW
mat_peptide
1..555
/product="unnamed"
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN
Query Match 100.0%; Score 561; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2e-118; Indels 0; Gaps 0;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTGGTAAAGTCAATGATACCTTACATGGGGTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTCAATGATACCTTACATGGGGTTCGCCGACCTCGTGGGGTACATT 60
Qy 61 CCGTCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCG 120
Db 61 CCGTCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCG 120
Qy 121 CTGAGGAGCGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTTCTATCTTC 180
Db 121 CTGAGGAGCGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTTCTATCTTC 180
Qy 181 CTCCTGGCTTGTCTGCTGCTGACCGTTCGACGCTTCGCGTTCGCGGAGAACAACTCTTCCGCG 240
Db 181 CTCCTGGCTTGTCTGCTGCTGACCGTTCGACGCTTCGCGTTCGCGGAGAACAACTCTTCCGCG 240
Qy 241 TCCGGGATGTACATGTACGAAAGAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 300
Db 241 TCCGGGATGTACATGTACGAAAGAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 300
Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTCGCGGAGAACAACTCTTCCGCG 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTCGCGGAGAACAACTCTTCCGCG 360
Qy 361 TGCTGGGTAGCGCTCACCCCAAGTCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 420
Db 361 TGCTGGGTAGCGCTCACCCCAAGTCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 420
Qy 421 ATACGAGCGCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 480
Db 421 ATACGAGCGCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 480
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAACCTGGTAATAG 561
Db 541 ATGATGATGAACCTGGTAATAG 561
RESULT 2
AR157337
LOCUS
DEFINITION
ACCESSION
VERSION
Sequence 23 from patent US 6245503.
AR157337
AR157337
AR157337.1 GI:16218270
linear PAT 17-OCT-2001
561 bp DNA
17-OCT-2001
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN
Query Match 100.0%; Score 561; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2e-118; Indels 0; Gaps 0;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTGGTAAAGTCAATGATACCTTACATGGGGTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTCAATGATACCTTACATGGGGTTCGCCGACCTCGTGGGGTACATT 60
Qy 61 CCGTCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCG 120
Db 61 CCGTCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCGACCTCGTGGGGTTCGCCG 120
Qy 121 CTGAGGAGCGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTTCTATCTTC 180
Db 121 CTGAGGAGCGCGTGAATATGCAACAGGGAATTTGCCCGGTTCCTTTCTATCTTC 180
Qy 181 CTCCTGGCTTGTCTGCTGCTGACCGTTCGACGCTTCGCGTTCGCGGAGAACAACTCTTCCGCG 240
Db 181 CTCCTGGCTTGTCTGCTGCTGACCGTTCGACGCTTCGCGTTCGCGGAGAACAACTCTTCCGCG 240
Qy 241 TCCGGGATGTACATGTACGAAAGAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 300
Db 241 TCCGGGATGTACATGTACGAAAGAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 300
Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTCGCGGAGAACAACTCTTCCGCG 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCGTTCGCGGAGAACAACTCTTCCGCG 360
Qy 361 TGCTGGGTAGCGCTCACCCCAAGTCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 420
Db 361 TGCTGGGTAGCGCTCACCCCAAGTCTGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 420
Qy 421 ATACGAGCGCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 480
Db 421 ATACGAGCGCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 480
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTTCACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAACCTGGTAATAG 561
Db 541 ATGATGATGAACCTGGTAATAG 561
RESULT 3
AX452772
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Sequence 23 from Patent EP1211315.
AX452772
AX452772.1 GI:21712457
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.

TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 23 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES Location/Qualifiers
source
1..561
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
CDS
1..558
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38073.1"
/db_xref="GI:21712458"
/translation="MLGKVIDLTTCGFADLVGYIPLVGAPLGGAAALAHGVRLVLEDD
VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNRVNSGMHYHTNDCNSSSIYVEAADM
INHTPGCVPCVRNNSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSYPGHITGHRMAMDMWNW"
mat_peptide
1..555
/product="unnamed"
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN
Query Match 100.0%; Score 561; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2e-118;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCGATACCCCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGTCAATCGATACCCCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGCGCTCGGGTT 120
DB 61 CGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGCGCTCGGGTT 120
QY 121 CTGGAGCAGCGCTGAACATGATCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGCAGCGCTGAACATGATCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTGTGACCGTTTCAGCTTCGCGTTTATGAAGTGGCAACGTTG 240
DB 181 CTCTGGCTTTGCTGCTGTGACCGTTTCAGCTTCGCGTTTATGAAGTGGCAACGTTG 240
QY 241 TCGGGATGATACATGTCAGAACGACTGCTCOACTCAAGCATTTGTGTATGAGGAGCG 300
DB 241 TCGGGATGATACATGTCAGAACGACTGCTCOACTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGCTGCTGCTGGGAGAACACTCTTCCCGC 360
DB 301 GACATGATCATGACACCCCGGGTGGCTGCTGCTGGGAGAACACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAGCGCCAGCGTCCCGCAGCACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAGCGCCAGCGTCCCGCAGCACA 420
QY 421 ATACGAGCGCAGTGTGATTCGAGTTCAGCATTCGCGCTCGCGGATGAGCGGTG 480
DB 421 ATACGAGCGCAGTGTGATTCGAGTTCAGCATTCGCGCTCGCGGATGAGCGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTTCAGGTATGGGTGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTTCAGGTATGGGTGGAT 540
QY 541 ATGATGATGAACCTGGTAATAG 561
DB 541 ATGATGATGAACCTGGTAATAG 561

RESULT 4
LOCUS AX685024
DEFINITION Sequence 23 from Patent WO0205548.
ACCESSION AX685024
VERSION AX685024.1 GI:29371429

KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F. and Buysse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: WO 02055548-A 23 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES Location/Qualifiers
source
1..561
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
CDS
1..558
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD8526.1"
/db_xref="GI:29371430"
/translation="MLGKVIDLTTCGFADLVGYIPLVGAPLGGAAALAHGVRLVLEDD
VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNRVNSGMHYHTNDCNSSSIYVEAADM
INHTPGCVPCVRNNSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSYPGHITGHRMAMDMWNW"
mat_peptide
1..555
/product="unnamed"
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN
Query Match 100.0%; Score 561; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2e-118;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCGATACCCCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGTCAATCGATACCCCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGCGCTCGGGTT 120
DB 61 CGCTCGTGGGGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATGCGCTCGGGTT 120
QY 121 CTGGAGCAGCGCTGAACATGATCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGCAGCGCTGAACATGATCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTGTGACCGTTTCAGCTTCGCGTTTATGAAGTGGCAACGTTG 240
DB 181 CTCTGGCTTTGCTGCTGTGACCGTTTCAGCTTCGCGTTTATGAAGTGGCAACGTTG 240
QY 241 TCGGGATGATACATGTCAGAACGACTGCTCOACTCAAGCATTTGTGTATGAGGAGCG 300
DB 241 TCGGGATGATACATGTCAGAACGACTGCTCOACTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGCTGCTGCTGGGAGAACACTCTTCCCGC 360
DB 301 GACATGATCATGACACCCCGGGTGGCTGCTGCTGGGAGAACACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAGCGCCAGCGTCCCGCAGCACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTAGGAGCGCCAGCGTCCCGCAGCACA 420
QY 421 ATACGAGCGCAGTGTGATTCGAGTTCAGCATTCGCGCTCGCGGATGAGCGGTG 480
DB 421 ATACGAGCGCAGTGTGATTCGAGTTCAGCATTCGCGCTCGCGGATGAGCGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTTCAGGTATGGGTGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTTCAGGTATGGGTGGAT 540
QY 541 ATGATGATGAACCTGGTAATAG 561
DB 541 ATGATGATGAACCTGGTAATAG 561

Db 421 ATACGAGCCACGTCGATTCCAGCTGTTACCACTCTCGCTCGCGGATGAGACGGTG 480

Qy 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGTCACCGTATGCTTGGGAT 540

Db 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGTCACCGTATGCTTGGGAT 540

Qy 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 6

ARL57338

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

556; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy 1 ATGTTGGGTAAAGTTCATCGATACGCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAAGTTCATCGATACGCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

Qy 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGCGCTTCGCCGACCTCGTGGGTACATT 120

Db 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGCGCTTCGCCGACCTCGTGGGTACATT 120

Qy 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Qy 181 CTCTTGGCTTTGCTGTCCTGTCGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGCTG 240

Db 181 CTCTTGGCTTTGCTGTCCTGTCGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGCTG 240

Qy 241 TCCGGGATGTACCATGTACGAAACGACCTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTACGAAACGACCTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Qy 301 GACATGATCATGACACACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 360

Db 301 GACATGATCATGACACACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 360

Qy 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 420

Db 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 420

Qy 421 ATAGACGCCACGCTGATTCACCGCTTCCAGCTTCCAGCTTCGCTTCGCGAGAACCACTCTTCCGCG 480

Db 421 ATAGACGCCACGCTGATTCACCGCTTCCAGCTTCCAGCTTCGCTTCGCGAGAACCACTCTTCCGCG 480

Qy 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGTCACCGTATGCTTGGGAT 540

Db 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGTCACCGTATGCTTGGGAT 540

Qy 541 ATGATGATGAACCTGGT 556

Db A48687

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

mat_peptide

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

556; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy 1 ATGTTGGGTAAAGTTCATCGATACGCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAAGTTCATCGATACGCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

Qy 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGCGCTTCGCCGACCTCGTGGGTACATT 120

Db 61 CCGCTCGTCGGCGCCCTAGGGGGCGCTGCCAGGCGCTTCGCCGACCTCGTGGGTACATT 120

Qy 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Qy 181 CTCTTGGCTTTGCTGTCCTGTCGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGCTG 240

Db 181 CTCTTGGCTTTGCTGTCCTGTCGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGCTG 240

Qy 241 TCCGGGATGTACCATGTACGAAACGACCTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTACGAAACGACCTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Qy 301 GACATGATCATGACACACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 360

Db 301 GACATGATCATGACACACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 360

Qy 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 420

Db 361 TGCTGGGTAGCGTCAACCCCGGTCGCTCGCTTCGCTTCGCGAGAACCACTCTTCCGCG 420

Qy 421 ATAGGAGCCACGCTCGATTCCAGCTTCCAGCTTCCAGCTTCGCTTCGCGAGAACCACTCTTCCGCG 480

```
Db 541 ATGATGATGAACCTGGT 556
|||||
RESULT 7
AX452774
LOCUS AX452774 606 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 25 from Patent EP1211315.
ACCESSION AX452774
VERSION AX452774.1 GI:21712459
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 25 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..606
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..603
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38074.1"
/db_xref="GI:21712460"
/translation="MLGKVIDLTTCGFADLVGYIPLVGAPLGGARALAHGVRLDGG
VNYATGNLPGCSFSIFLLALLSLCTVPASAYEVNRVSGMYHVTNDCNSGSIYVEADM
IMHTGCVPCVRENSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QPCNCSIYPGHITGHRMAMWMMNSPTTALVVSQLLRL"
mat_peptide
1..600
/product="unnamed"
BASE COUNT 109 a 193 c 167 g 137 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCATACCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATCATACCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTCGCGCATGCGCTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTCGCGCATGCGCTCCGGTT 120
QY 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180
Db 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGTCCTGTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACG 240
Db 181 CTCCTGGCTTTGCTGTCCTGTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACG 240
QY 241 TCGGGATGATCAATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300
Db 241 TCGGGATGATCAATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCAACCCCGGTCGTCCTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCAACCCCGGTCGTCCTCGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTCAACCCCGCAGCTCGAGCTAGGAAGCCAGCGTCCGCCACGACA 420
Db 361 TGTGGGTAGCGCTCAACCCCGCAGCTCGAGCTAGGAAGCCAGCGTCCGCCACGACA 420
QY 421 ATACGACGCGCATGTCAGTTCACAGCTGTTCAACATCTGCGCTCGCGCATGACGGTG 480
|||||
```

```
Db 421 ATACGACGCGCATGTCAGTTCACAGCTGTTCAACATCTCGCTCGCGCATGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCTCGGAT 540
|||||
Db 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATGGCTCGGAT 540
QY 541 ATGATGATGAACCTGGT 556
|||||
Db 541 ATGATGATGAACCTGGT 556
|||||
RESULT 8
AX685026
LOCUS AX685026 606 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 25 from Patent WO0205548.
ACCESSION AX685026
VERSION AX685026.1 GI:29371431
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1
AUTHORS Maertens,G., Bosman,F. and Buyse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 25 18-JUL-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..606
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..603
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86527.1"
/db_xref="GI:29371432"
/translation="MLGKVIDLTTCGFADLVGYIPLVGAPLGGARALAHGVRLDGG
VNYATGNLPGCSFSIFLLALLSLCTVPASAYEVNRVSGMYHVTNDCNSGSIYVEADM
IMHTGCVPCVRENSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QPCNCSIYPGHITGHRMAMWMMNSPTTALVVSQLLRL"
mat_peptide
1..600
/product="unnamed"
BASE COUNT 109 a 193 c 167 g 137 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCATACCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATCATACCTTATACGCGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTCGCGCATGCGCTCCGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTCGCGCATGCGCTCCGGTT 120
QY 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180
Db 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCGGTTCCTTTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGTCCTGTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACG 240
Db 181 CTCCTGGCTTTGCTGTCCTGTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACG 240
QY 241 TCGGGATGATCAATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300
Db 241 TCGGGATGATCAATGTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCAACCCCGGTCGTCCTCGGAGAACAACTCTTCCCGC 360
|||||
```

Db 301 GACATGATCATCCACACCCCGGGTGGCTGGCTGGCTGGCGGAGAACAACTCTTCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAAAGCGCAGGCTCCCAACGACGA 420

Db 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAAAGCGCAGGCTCCCAACGACGA 420

QY 421 ATACGACGCGCATGCTGATCCAGCTGTTCCACCTTCGCTCGCGGCATGAGACGGTG 480

Db 421 ATACGACGCGCATGCTGATCCAGCTGTTCCACCTTCGCTCGCGGCATGAGACGGTG 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACGGGTACCGGTATGGCTGGGAT 540

Db 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACGGGTACCGGTATGGGAT 540

QY 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 9

LOCUS A48689 636 bp DNA linear PAT 07-MAR-1997

DEFINITION Sequence 27 from Patent WO9604385.

ACCESSION A48689

VERSION A48689.1 GI:2302402

KEYWORDS unidentified

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 636)

AUTHORS Maertens, G., Bosman, F., De, M. G. and Buyse, M.

TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

JOURNAL Patent: WO 9604385-A 27 15-FEB-1996;

COMMENT INNOGENETICS NV (BE)

Other publication CA 2172273 960215

Other publication AU 3382495 960304.

FEATURES

source

1..636

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

1..633

/note="unnamed protein product"

/codon_start=1

/protein_id="CAA03137.1"

/db_xref="GI:2302403"

/translation="MLGKVIIDLTGCFADLVGYIPLVGLGGAARALAHGVRVLEDG VNYATGNLPGCSFIFLLALLSLCTVPASAYEVNVSVMVHTNDNSNIVVEAADM IMHTGPCVPCVRENNSRCWALPTTLAARNASVPTTIRRHVDSQLFTISPRHEIV ODCNCSIVPGHITGHRMAWDMNWSPTTALVVSQULRVITGRRHHHHH"

mat_peptide 1..630

/product="unnamed"

BASE COUNT 119 a 203 c 174 g 140 t

ORIGIN

Query Match 99.1%; Score 556; DB 6; Length 636;

Best Local Similarity 100.0%; Pred. No. 2.8e-117;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCAATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60

Db 1 ATGTTGGGTAAAGTCAATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60

QY 61 CCCTCGTCCGGCCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

Db 61 CCCTCGTCCGGCCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGAGAGACGGCGTGAACACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTC 180

Db 121 CTGAGAGACGGCGTGAACACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTC 180

QY 181 CTCTTGGCTTTGGTGTCTCTCTGACCGTTCAGCTTCAGTTCAGTTCAGTTCAGTTCAGTTC 240

Db 181 CTCTTGGCTTTGGTGTCTCTCTGACCGTTCAGCTTCAGTTCAGTTCAGTTCAGTTCAGTTC 240

QY 241 TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Db 181 CTCTTGGCTTTGGTGTCTCTCTGACCGTTCAGCTTCAGTTCAGTTCAGTTCAGTTCAGTTC 240

QY 241 TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

QY 301 GACATGATCATGACACACCCCGGGTGGCTGGCTGGCTGGCGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGCTGGCTGGCTGGCGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAAAGCGCAGGCTCCCAACGACGA 420

Db 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAAAGCGCAGGCTCCCAACGACGA 420

QY 421 ATACGACGCGCATGCTGATTCGCTGATTCGCTGATTCGCTGATTCGCTGATTCGCTGATTC 480

Db 421 ATACGACGCGCATGCTGATTCGCTGATTCGCTGATTCGCTGATTCGCTGATTCGCTGATTC 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

Db 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

QY 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 10

LOCUS ARI57339 636 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 27 from patent US 6245503.

ACCESSION ARI57339

VERSION ARI57339.1 GI:16218273

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 636)

AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.

TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use

JOURNAL Patent: US 6245503-A 27 12-JUN-2001;

FEATURES

Location/Qualifiers

source 1..636

/organism="unknown"

BASE COUNT 119 a 203 c 174 g 140 t

ORIGIN

Query Match 99.1%; Score 556; DB 6; Length 636;

Best Local Similarity 100.0%; Pred. No. 2.8e-117;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCAATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60

Db 1 ATGTTGGGTAAAGTCAATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60

QY 61 CCCTCGTCCGGCCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

Db 61 CCCTCGTCCGGCCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGAGAGACGGCGTGAACACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTC 180

Db 121 CTGAGAGACGGCGTGAACACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATCTTC 180

QY 181 CTCTTGGCTTTGGTGTCTCTCTGACCGTTCAGCTTCAGTTCAGTTCAGTTCAGTTCAGTTC 240

Db 181 CTCTTGGCTTTGGTGTCTCTCTGACCGTTCAGCTTCAGTTCAGTTCAGTTCAGTTCAGTTC 240

QY 241 TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

Db 241 TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCG 300

```
QY 301 GACATGATCATGCACACCCCGGCTGCTGCCCTGCTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGCTGCTGCCCTGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGTAGCGCTCACACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
Db 361 TCGTGGTAGCGCTCACACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
QY 421 ATACGAGCCGCGTCAATTCACCGCTGTTCAACCATCTCGCTCGCGGAGATGAGCGGTG 480
Db 421 ATACGAGCCGCGTCAATTCACCGCTGTTCAACCATCTCGCTCGCGGAGATGAGCGGTG 480
QY 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGGTGGGAT 540
Db 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGGTGGGAT 540
QY 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 11
AX452776
LOCUS AX452776 636 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 27 from Patent EP1211315.
ACCESSION AX452776
VERSION AX452776.1 GI:21712461
KEYWORDS Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 27 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..636
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38075.1"
/db_xref="GI:21712462"
VNYATGNLPGCSFSIFLLALLSCLTVPASAEVRNVSGMHTVNDSCNSSIIVVEADM
IMHTPGCPVCRVNSSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QPCNCSIYPGHITGHRMADMMNWSPTTALVVSQLLRIVIBGRHHHHH
mat_peptide
1..630
/product="unnamed"
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGTCCGGTT 120
Db 61 CCGCTCGTGGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGTCCGGTT 120
QY 121 CTGGAGGAGCGGTGTAACATATGCAACAGGGAATTTGCCGGTGTGCTTCTCTATCTTC 180
Db 121 CTGGAGGAGCGGTGTAACATATGCAACAGGGAATTTGCCGGTGTGCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACAGTG 240
```

```
Db 181 CTCTTGGCTTTGCTGTCCTGTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACAGTG 240
QY 241 TCGGAGTGTACCATGTACGAACGACTGCTCCAATCTCAAGCATTTGTGTATGAGGAGCG 300
Db 241 TCGGAGTGTACCATGTACGAACGACTGCTCCAATCTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGCTGCTGCCCTGCTCGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGCTGCTGCCCTGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGTAGCGCTCACACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
Db 361 TCGTGGTAGCGCTCACACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCCACACGACA 420
QY 421 ATACGAGCCGCGTCAATTCACCGCTGTTCAACCATCTCGCTCGCGGAGATGAGCGGTG 480
Db 421 ATACGAGCCGCGTCAATTCACCGCTGTTCAACCATCTCGCTCGCGGAGATGAGCGGTG 480
QY 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGGTGGGAT 540
Db 481 CAGGACTGCAATTCCTCAATCTATCCCGCCACATAACGGGTACCGTATGGGTGGGAT 540
QY 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 12
AX685028
LOCUS AX685028 636 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 27 from Patent WO0205548.
ACCESSION AX685028
VERSION AX685028.1 GI:29371433
KEYWORDS Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 27 18-JUL-2002;
Innogenetics N.V. (BE)
FEATURES
Location/Qualifiers
source
1..636
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86528.1"
/db_xref="GI:29371434"
VNYATGNLPGCSFSIFLLALLSCLTVPASAEVRNVSGMHTVNDSCNSSIIVVEADM
IMHTPGCPVCRVNSSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QPCNCSIYPGHITGHRMADMMNWSPTTALVVSQLLRIVIBGRHHHHH
mat_peptide
1..630
/product="unnamed"
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN
Query Match 99.1%; Score 556; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCGCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGTCCGGTT 120
```



```
Db 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CGCTCGTGGGCGCCCTAGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGT 120
Db 61 CGCTCGTGGGCGCCCTAGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGT 120
Qy 121 CTGGAGGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Db 121 CTGGAGGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTCCCAACGTTG 240
Db 181 CTCTGGCTTTGCTGCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTCCCAACGTTG 240
Qy 241 TCCGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 300
Db 241 TCCGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 300
Qy 421 ATACGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Db 421 ATACGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Qy 301 GACATGATCATGACACACCCCGGTCGCTGCGGTCGCTGCGGAGAACCACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGTCGCTGCGGTCGCTGCGGAGAACCACTCTTCCCGC 360
Qy 361 TCGTGGTAGCGTCAACCTGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 300
Db 361 TCGTGGTAGCGTCAACCTGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 300
Qy 421 ATACGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Db 421 ATACGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTTCACCGTATGGCTTGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTTCACCGTATGGCTTGGAT 540
Qy 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 15
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
IMMUNOGENETICS N.V. (BE)
FEATURES
source
1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD38072.1"
/db_xref="GI:21712456"
/translation="MLGKVIDITLTCGFADLVGYIPLVGAPLGAARALAHGVRVLDEG
VYATGNLPGCSFISIFLLALLSCLTPASAYEVNRYSGMYHVNDCNSISIVYEADM
IMHTPGCPVCRNNSRCWALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGCHITGHRMAMWMMNWSPTTALVVSQLLRIPQAVVDVMAHGVLAGL
AYYSVMGNKVLIVMLFAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
```

ORIGIN

```
Query Match 99.1%; Score 556; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.8e-117;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAAGTCAATCGATACCCCTTACATGGGCTTCGCCGACCTCGTGGGTACATT 60
Qy 61 CGCTCGTGGGCGCCCTAGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGT 120
Db 61 CGCTCGTGGGCGCCCTAGGGGCGGTGCCAGGGCCCTGCGCATGCGTCCGGGT 120
Qy 121 CTGGAGGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Db 121 CTGGAGGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTCCCAACGTTG 240
Db 181 CTCTGGCTTTGCTGCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTCCCAACGTTG 240
Qy 241 TCCGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 300
Db 241 TCCGGATGTACATGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 300
Qy 301 GACATGATCATGACACACCCCGGTCGCTGCGGTCGCTGCGGAGAACCACTCTTCCCGC 360
Db 301 GACATGATCATGACACACCCCGGTCGCTGCGGTCGCTGCGGAGAACCACTCTTCCCGC 360
Qy 361 TCGTGGTAGCGTCAACCTGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 420
Db 361 TCGTGGTAGCGTCAACCTGTCAGAACGACTGCTCCAACTCAAGCATTTGTTATGAGGAGCG 420
Qy 421 ATACGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 480
Db 421 ATACGAGCGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTCTATCTTC 480
Qy 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTTCACCGTATGGCTTGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTTCACCGTATGGCTTGGAT 540
Qy 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556
```

Search completed: December 20, 2003, 02:02:02
Job time : 2126.1 Secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 155.093 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-23

Perfect score: 561

Sequence: 1 ATGTTGGTAAAGTCATCGA.....TCATGATGAACGTGTAATAG 561

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	561	100.0	561	17	AAT12962 HCV E1 construct H
2	561	100.0	561	24	AAL48926 Hepatitis C virus
3	556	99.1	606	17	AAT12963 HCV E1 construct H
4	556	99.1	606	24	AAL48927 Hepatitis C virus
5	556	99.1	636	17	AAT12964 HCV E1 construct H
6	556	99.1	636	24	AAL48928 Hepatitis C virus
7	556	99.1	723	17	AAT12961 HCV E1 construct H
8	556	99.1	723	24	AAL48925 Hepatitis C virus

9	474	84.5	795	17	AAT12705 HCV E1 construct H
10	474	84.5	795	24	AAL48914 Hepatitis C virus
11	471	84.0	2082	24	AAL48939 Hepatitis C virus
12	471	84.0	2086	17	AAT12973 HCV E1 construct H
13	471	84.0	2433	17	AAT12974 HCV E1 construct H
14	461.4	82.2	633	17	AAT12706 HCV E1 construct H
15	461.4	82.2	633	24	AAL48915 Hepatitis C virus
16	460	82.0	2434	24	AAL48940 HCV E1 construct H
17	454	80.9	636	17	AAT12709 Hepatitis C virus
18	454	80.9	636	24	AAL48918 HCV core protein c
19	423.6	75.5	673	19	AAV42305 Hepatitis C virus
20	420.2	74.9	9605	24	ABK91411 Hepatitis C virus
21	420.2	74.9	9605	24	ABK91424 Hepatitis C virus
22	420.2	74.9	9605	24	ABK91425 Hepatitis C virus
23	420.2	74.9	9605	24	ABK91426 Hepatitis C virus
24	420.2	74.9	9605	24	ABK91428 Hepatitis C virus
25	420.2	74.9	9605	24	ABK91429 Hepatitis C virus
26	420.2	74.9	9605	24	ABK91430 Hepatitis C virus
27	420.2	74.9	9605	24	ABK91431 Hepatitis C virus
28	420.2	74.9	9605	24	ABK91432 Hepatitis C virus
29	420.2	74.9	9605	24	ABK91433 Hepatitis C virus
30	420.2	74.9	9605	24	ABK91433 Hepatitis C virus
31	420.2	74.9	9608	24	ABK91427 Hepatitis C virus
32	420.2	74.9	11062	24	ABK91427 Hepatitis C virus
33	420.2	74.9	11076	21	AA98965 Hepatitis C virus
34	418.6	74.6	1562	19	AAV60672 Fragment #5 isolat
35	418.6	74.6	1953	25	AAL55222 Plasmid pIDKE2 DNA
36	418.6	74.6	2187	19	ABA03491 Cuticle protein 1
37	418.6	74.6	2540	14	AAQ43889 NANB hepatitis vir
38	418.6	74.6	2540	15	AAQ63753 NANBH genomic fra
39	418.6	74.6	2829	19	AAV60673 Fragment #6 isolat
40	418.2	74.5	1270	19	AAV60668 Fragment #1 isolat
41	417	74.3	1880	13	AAQ24467 NANB hepatitis vir
42	415.4	74.0	2540	13	AAQ29628 Hepatitis C virus
43	413.8	73.8	1863	12	AAQ15363 Fragment of NANB h
44	413.4	73.7	1037	15	AAQ58449 Hepatitis C virus
45	413.4	73.7	9609	24	AAQ33038 HCV-S1 full-length

ALIGNMENTS

RESULT 1
AAT12962
ID AAT12962 standard; DNA; 561 BP.

AC AAT12962;

XX 24-SEP-1996 (first entry)

XX HCV E1 construct HCCI38.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.

OS Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

XX Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

PT proteins - in presence of di:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
PS
XX
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX
SQ Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;

Query Match 100.0%; Score 561; DB 17; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.5e-140;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCAACCCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCAACCCTCGTGGGTACATT 60

QY 61 CCGCTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGCTCCGGTT 120
DB 61 CCGCTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGCTCCGGTT 120

QY 121 CTGAGGACGCGGTGAATATGCAACAGAGGAATTTGCCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGAGGACGCGGTGAATATGCAACAGAGGAATTTGCCGGTTGCTCTTCTCTATCTTC 180

QY 181 CTCCTTGGCTTTGCTGCTGTGTCCTGTCGACCGTTCCAGCTTCCAGCTTCCAGCTG 240
DB 181 CTCCTTGGCTTTGCTGCTGTGTCCTGTCGACCGTTCCAGCTTCCAGCTTCCAGCTG 240

QY 241 TCCGGGATGTACCATGTTCAGACGACCTGCTCCAACTCAAGCATTTGTATGAGCGAGCG 300
DB 241 TCCGGGATGTACCATGTTCAGACGACCTGCTCCAACTCAAGCATTTGTATGAGCGAGCG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTGCCCTCGGTTGGGAGAACACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTGCCCTCGGTTGGGAGAACACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCGAGCTCGCAGCTAGGAGCGCCAGCGTCCCGGATGAGCGGTG 420
DB 361 TGCTGGGTAGCGCTCACCCCGAGCTCGCAGCTAGGAGCGCCAGCGTCCCGGATGAGCGGTG 420

QY 421 ATACGACGCGCATGTCGATTTCCAGCTGTTTCAACATCTCGCTCGCGGATGAGCGGTG 480
DB 421 ATACGACGCGCATGTCGATTTCCAGCTGTTTCAACATCTCGCTCGCGGATGAGCGGTG 480

QY 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCCATTAACGGGTTCACCGTATGCTTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCGGCCCATTAACGGGTTCACCGTATGCTTTGGGAT 540

QY 541 ATGATGATGAACCTGGTAAATAG 561
DB 541 ATGATGATGAACCTGGTAAATAG 561

AA148926
ID AAL48926 standard; DNA; 561 BP.
XX
AC AAL48926;
XX
DT 24-OCT-2002 (first entry)
XX
XX Hepatitis C virus clone HCCI38 E1 protein coding sequence.
DE
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW viricide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
XX WO200255548-A2.
XX
PD 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
XX
XX 30-AUG-2001; 2001US-315768P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buyse M;
XX
XX WPI: 2002-599657/64.
XX
XX P-PSDB; AAO18667.
XX
XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection.
XX
XX Example 2; Page 175-176; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;
SQ

Query Match 100.0%; Score 561; DB 24; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.5e-140;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCAACCCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCAACCCTCGTGGGTACATT 60

QY 61 CCGCTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGCTCCGGTT 120
DB 61 CCGCTCGTCGCGCGCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGCTCCGGTT 120

QY 121 CTGAGGACGCGGTGAATATGCAACAGAGGAATTTGCCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGAGGACGCGGTGAATATGCAACAGAGGAATTTGCCGGTTGCTCTTCTCTATCTTC 180

QY 181 CTCCTTGGCTTTGCTGCTGTGTCCTGTCGACCGTTCCAGCTTCCAGCTTCCAGCTG 240
DB 181 CTCCTTGGCTTTGCTGCTGTGTCCTGTCGACCGTTCCAGCTTCCAGCTTCCAGCTG 240

QY 241 TCCGGGATGTACCATGTTCAGACGACCTGCTCCAACTCAAGCATTTGTATGAGCGAGCG 300

PF 11-JAN-2002; 2002WO-EP00219.
XX 11-JAN-2001; 2001US-260699P.
PR 30-AUG-2001; 2001US-315768P.
XX (INNO-) INNOGENETICS NV.
XX Maertens G, Bosman F, Buyse M;
XX WPI; 2002-599657/64.
DR P-PSDB; AAO18668.
XX New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX Example 2; Page 177-178; 243pp; English.
XX The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV) specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 606 BP; 109 A; 193 C; 167 G; 137 T; 0 other;
Query Match 99.1%; Score 556; DB 24; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.2e-138;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCTGCCAGGCGCTGGCGCATGGCTCCGGT 120
DB 61 CCGCTCGTGGGCGCCCTAGGGGGCTGCCAGGCGCTGGCGCATGGCTCCGGT 120
QY 121 CTGAGGACGCGGTGAATATGCAACAGGGAATTTGCCCGTTGCTCTTCTATCTTC 180
DB 121 CTGAGGACGCGGTGAATATGCAACAGGGAATTTGCCCGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGTGCTGACCGTTCCAGTTCGGTTTATGAAGTGGCAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTGTGCTGACCGTTCCAGTTCGGTTTATGAAGTGGCAAGTG 240
QY 241 TCCGGGATGATACCATGTCAGAACGACGTCCCACTCAAGCATTTGTATGAGCGAGCG 300
DB 241 TCCGGGATGATACCATGTCAGAACGACGTCCCACTCAAGCATTTGTATGAGCGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTGCGCTGCGTTCGGGAGAAACAATCTTCGCG 360
DB 301 GACATGATCATGCACACCCCGGGTGCCTGCGCTGCGTTCGGGAGAAACAATCTTCGCG 360
QY 361 TCGTGGGTAGCGCTCACCCCGACGCTCGAGTAGAAGCGCGGTCCCGACACGACA 420
DB 361 TCGTGGGTAGCGCTCACCCCGACGCTCGAGTAGAAGCGCGGTCCCGACACGACA 420
QY 421 ATACGACGCGCATGATGATCCAGAGTGTTCACCATCTCGCTTCGCGCATGAGCGGTG 480
DB 421 ATACGACGCGCATGATGATCCAGAGTGTTCACCATCTCGCTTCGCGCATGAGCGGTG 480
QY 481 CAGGACTGCAATGTCTCAATATATCCCGGCGCATTAACGGGTACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATGTCTCAATATATCCCGGCGCATTAACGGGTACCGTATGGCTTGGAT 540

QY 541 ATGATGATGAAGTGGT 556
DB 541 ATGATGATGAAGTGGT 556
RESULT 5
AAT12964
ID AAT12964 standard; DNA; 636 BP.
XX AAT12964;
XX 24-SEP-1996 (first entry)
XX HCV E1 construct HCC140.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
SQ Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;
Query Match 99.1%; Score 556; DB 17; Length 636;
Best Local Similarity 100.0%; Pred. No. 1.2e-138;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTAGGGGGCTGCCAGGCGCTGGCGCATGGCTCCGGT 120

Query Match 99.1%; Score 556; DB 24; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.3e-138;


```
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTCAATCGATACCCCTTATGCGGCTTCCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGTCAATCGATACCCCTTATGCGGCTTCCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTT 120
DB 61 CCGCTCGTGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTT 120
QY 121 CTGGAGGCGGCTGAATCATGATCAACAGGGAATTTGCCCCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGCGGCTGAATCATGATCAACAGGGAATTTGCCCCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCAACGTTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCAACGTTG 240
QY 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 360
DB 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 360
QY 361 TGCTGGGTAGCGCTACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 420
DB 361 TGCTGGGTAGCGCTACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 420
QY 421 ATACGAGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 ATACGAGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATATAACGGGTCAACGGGTATGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGCATATAACGGGTCAACGGGTATGGGAT 540
QY 541 ATGATGATGAATGGT 556
DB 541 ATGATGATGAATGGT 556

RESULT 9
AAT12705
ID AAT12705 standard; DNA; 795 BP.
AC AAT12705;
XX
XX
XX 23-SEP-1996 (first entry)
DE HCV E1 construct 'HCC110A'.
XX
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
XX Hepatitis C virus.
XX
XX WO9604385-A2.
XX
XX 15-FEB-1996.
XX
XX 31-JUL-1995; 95WO-BP03031.
XX
XX 29-JUL-1994; 94EP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX
```

```
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
PS Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;
SQ
Query Match 84.5%; Score 474; DB 17; Length 795;
Best Local Similarity 88.5%; Pred. No. 1e-116; Mismatches 0; Indels 72; Gaps 1;
Matches 556; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGTCAATCGATACCCCTTATGCGGCTTCCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGTCAATCGATACCCCTTATGCGGCTTCCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTT 120
DB 61 CCGCTCGTGGCGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGTCCGGGTT 120
QY 121 CTGGAGGCGGCGTGAATCATGATCAACAGGGAATTTGCCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGCGGCGTGAATCATGATCAACAGGGAATTTGCCCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCAACGTTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCCAGCTTCCGCTTATGAAGTGGCAACGTTG 240
QY 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 360
DB 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 360
QY 361 TGCTGGGTAGCGCTACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 420
DB 361 TGCTGGGTAGCGCTACACCCCGGGTGGTGGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 420
QY 421 ATACGAGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 ATACGAGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 439 -----TCCAGCTGTTTACCATCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 468
DB 481 GGGGACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 469 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGGCACATAAGCGGTTCACCGT 528
DB 541 CATGAGACGGTGCAGGACTGCAATTCATATCCCGGCGGCACATAAGCGGTTCACCGT 600
QY 529 ATGGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
```

181	CTCTTGGCTTTGCTGCTCTGACCGGTTCCAGCTTCCGCTTATGAAGTGCACAACGTTG	240
QY		
181	CTCTTGGCTTTGCTGCTCTGACCGGTTCCAGCTTCCGCTTATGAAGTGCACAACGTTG	
Db		
241	TCCGGGATGTACCATGTCACGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGCAGCG	300
QY		
241	TCCGGGATGTACATGTCACGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGCAGCG	300
Db		
301	GACATGATCATGACACACCCCGGGTGCCTGCTCGTTGGGAGAACAACTCTTCCCGC	360
QY		
301	GACATGATCATGACACACCCCGGGTGCCTGCTCGTTGGGAGAACAACTCTTCCCGC	360
Db		
361	TGCTGGGTAGCGTCAACCCCGCAGCTCGCAGCTAGCAACGCGCAGCTCCCCACCAACGACA	420
QY		
361	TGCTGGGTAGCGTCAACCCCGCAGCTCGCAGCTAGCAACGCGCAGCTCCCCACCAACGACA	420
Db		
421	ATACGAGCCACGTCGAT	438
QY		
421	ATACGAGCCACGTCGATTTGCTGTTGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG	480
Db		
439	-----TCCAGCTGTTCCACATCTCCGCTCGCGG	468
QY		
481	GGGACCTCTCGGATCTGTCCTCCGCTCCACATCTTCGCGCTCGCGG	540
Db		
469	CATGAGCGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGT	528
QY		
541	CATGAGCGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGT	600
Db		
529	ATGGCTTGGGATATGATGAAGAACTGGT	556
QY		
601	ATGGCTTGGGATATGATGAAGAACTGGT	628
Db		

RESULT 11	
AA148939	standard; DNA; 2082 BP.
ID	AA148939
XX	
AC	AA148939;
XX	
DT	24-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.
XX	
KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	viridae; immunostimulant; vaccine; ds.
XX	
OS	Hepatitis C virus.
XX	
PN	WO200255548-A2.
XX	
PD	18-JUL-2002.
XX	
PF	11-JAN-2002; 2002WO-EP00219.
XX	
PR	11-JAN-2001; 2001US-260699P.
XX	
PR	30-AUG-2001; 2001US-315768P.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Maertens G, Bosman F, Buyse M;
XX	
DR	WPI; 2002-599657/64.
DR	P-PSDB; AAC18678.
XX	
PT	New therapeutic vaccine compositions comprising at least one purified
PT	recombinant hepatitis C virus (HCV) single or specific oligomeric
PT	recombinant envelope protein E1 or E2, useful for immunizing humans
PT	from HCV infection
XX	
PS	Example 2; Page 206-209; 243pp; English.
XX	
CC	The present invention relates to new therapeutic vaccine compositions fo
CC	inducing hepatitis C virus (HCV)-specific antibodies, comprising a

Db 601 ATGGCTTGGGATGATGATGAACGTGT 628

|||||

RESULT 10
 AAL48914
 ID AAL48914 standard; DNA; 795 BP.
 XX
 AC AAL48914;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus clone HCC110A E1 protein coding sequence.
 XX
 KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 OS
 OS Hepatitis C virus.
 PN WO200255548-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EP00219.
 XX
 PR 11-JAN-2001; 2001US-260699P.
 PR 30-AUG-2001; 2001US-315768P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Maertens G, Bosman F, Buyse M;
 DR WPI; 2002-599657/64.
 DR P-PSDB; RAO18661.
 XX
 PT New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection -
 XX
 XX Example 2; Page 161-162; 243pp: English.
 XX
 CC The present invention relates to new therapeutic vaccine compositions for
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX
 XX Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;
 SQ

Query Match 84.5%; Score 474; DB 24; Length 795;
 Best Local Similarity 88.5%; P-red. No. 1e-116;
 Matches 536; Conservative 0; Mismatches 0; Indels 72; Gaps 1

QY 1 ATCTTGGGTAAGTCTATCGATACCCCTTACATGGCGCTTCGCCGACCTCGTGGGTACATT 60
 DB 1 ATGTTGGGTAAGTCTATCGATACCCCTTACATGGCGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCCTCGTGGCGCCCCCTTAGGGGCGGTGCCAGGGGCCCTGGCGCATGGCGTCCGGGTT 120
 DB 61 CCGCTCGTGGCGCCCCCTTAGGGGCGGTGCCAGGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGAGGACGGCGTGAACTATGCAACAGGAAATTGCCGGTGTGCTTTCTCTATCTTC 180
 DB 121 CTGAGACGGCGTGAACTATGCAACAGGAAATTGCCGGTGTGCTTTCTCTATCTTC 180

CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 2086 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

Query Match 84.0%; Score 471; DB 24; Length 2082;
Best Local Similarity 88.5%; Pred. No. 8.4e-116;
Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTTCATCGATACCCCTTACATGCGGGTTCGCCGACCTCGTGGGGTACATTCCG 63
Db |||||
QY 4 TTGGGTAAGTTCATCGATACCCCTTACATGCGGGTTCGCCGACCTCGTGGGGTACATTCCG 63
Db |||||
QY 64 CTGCTGGCGCCCCCTAGGGGGCGTGCAGGGCCCTGCGCATGCGTCCGGGTTCG 123
Db |||||
QY 64 CTGCTGGCGCCCCCTAGGGGGCGTGCAGGGCCCTGCGCATGCGTCCGGGTTCG 123
Db |||||
QY 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTCTATCTTCCTC 183
Db |||||
QY 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTCTATCTTCCTC 183
Db |||||
QY 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCGCAACGTGTC 243
Db |||||
QY 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCGCAACGTGTC 243
Db |||||
QY 244 GGGATGTACATGTCAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
Db |||||
QY 304 ATGATATGACACACCCCGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db |||||
QY 304 ATGATATGACACACCCCGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db |||||
QY 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCGACACGACAATA 423
Db |||||
QY 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCGACACGACAATA 423
Db |||||
QY 424 CGAGCCACGTCGAT----- 438
Db |||||
QY 424 CGAGCCACGTCGATTTGCTGTTGGGGCGGCTCTTCTGTTCCGCTATGATGTTGGG 483
QY 439 -----TCCAGCTGTTCAACATCTGCTCGCTCGCGGCTAT 471
Db |||||
QY 484 GACCTCTGCGGATCTGTCTTCTCTGCTCTCCAGCTGTTCAACATCTGCTCGCGGCTAT 543
QY 472 GAGACGCTGACGAGCTGCAATTCCTCAATCTATCCGGCCACATACGGGTACCGGTATG 531
Db |||||
QY 544 GAGACGCTGACGAGCTGCAATTCCTCAATCTATCCGGCCACATACGGGTACCGGTATG 603
QY 532 GCTTGGGATATGATGATGAACTGTT 556
Db |||||
QY 604 GCTTGGGATATGATGATGAACTGTT 628

RESULT 12
AA12973
ID AA12973 standard; DNA; 2086 BP.
XX
AC AA12973;
XX
XX 24-SBP-1996 (first entry)
DT
XX
DE HCV E1 construct HCC165.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95MO-EP03031.
XX
PR 29-JUL-1994; 94BP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
PA Bosman P, Buyse M, De Martynoff G, Maertens G;
PI WPI; 1996-129401/13.
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di: sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AA12704-T12709 and AA12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
SQ Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;

Query Match 84.0%; Score 471; DB 17; Length 2086;
Best Local Similarity 88.5%; Pred. No. 8.4e-116;
Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTTCATCGATACCCCTTACATGCGGGTTCGCCGACCTCGTGGGGTACATTCCG 63
Db |||||
QY 4 TTGGGTAAGTTCATCGATACCCCTTACATGCGGGTTCGCCGACCTCGTGGGGTACATTCCG 63
Db |||||
QY 64 CTGCTGGCGCCCCCTAGGGGGCGTTCGCCGAGCGCTGCGCATGCGTCCGGGTTCG 123
Db |||||
QY 64 CTGCTGGCGCCCCCTAGGGGGCGTTCGCCGAGCGCTGCGCATGCGTCCGGGTTCG 123
Db |||||
QY 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTCTATCTTCCTC 183
Db |||||
QY 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTCTATCTTCCTC 183
Db |||||
QY 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCGCAACGTGTC 243
Db |||||
QY 184 TTGGCTTTTGTCTGCTGTCTGACCGTTCCAGCTTCGCTTATGAAGTGCGCAACGTGTC 243
Db |||||
QY 244 GGGATGTACATGTCAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
Db |||||
QY 244 GGGATGTACATGTCAGAACGACTGTCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
Db |||||
QY 304 ATGATATGACACACCCCGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db |||||
QY 304 ATGATATGACACACCCCGGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db |||||
QY 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCGACACGACAATA 423
Db |||||
QY 364 TGGGTAGCGCTCACCCCGGCTCGAGCTAGGAAGCCAGCGTCCCGACACGACAATA 423
Db |||||
QY 424 CGAGCCACGTCGAT----- 438
Db |||||
QY 424 CGAGCCACGTCGATTTGCTGTTGGGGCGGCTCTTCTGTTCCGCTATGATGTTGGG 483
QY 439 -----TCCAGCTGTTCAACATCTGCTCGCTCGCGGCTAT 471
Db |||||
QY 484 GACCTCTGCGGATCTGTCTTCTCTGCTCTCCAGCTGTTCAACATCTGCTCGCGGCTAT 543
QY 472 GAGACGCTGACGAGCTGCAATTCCTCAATCTATCCGGCCACATACGGGTACCGGTATG 531
Db |||||
QY 544 GAGACGCTGACGAGCTGCAATTCCTCAATCTATCCGGCCACATACGGGTACCGGTATG 603
QY 532 GCTTGGGATATGATGATGAACTGTT 556
Db |||||
QY 604 GCTTGGGATATGATGATGAACTGTT 628

RESULT 12
AA12973
ID AA12973 standard; DNA; 2086 BP.
XX
AC AA12973;
XX
XX 24-SBP-1996 (first entry)
DT
XX
DE HCV E1 construct HCC165.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;

Query Match 84.0%; Score 471; DB 17; Length 2433;
Best Local Similarity 88.5%; Pred. No. 8.8e-116; Indels 72; Gaps 1;
Matches 553; Conservative 0; Mismatches 0;

4 TTGGGTAAAGTGCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCG 63
355 TTGGGTAAAGTGCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCG 414
64 CTCGTGCGCGCGCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 123
415 CTCGTGCGCGCGCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCGTTCGGGTCTG 474
124 GAGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTCTC 183
475 GAGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTCTC 534
184 TTGGCTTTGCTGCTCTGCTGACCCGTTCCAGCTTCGCTTATGAAGTGCAGCGTCTCC 243
535 TTGGCTTTGCTGCTCTGCTGACCCGTTCCAGCTTCGCTTATGAAGTGCAGCGTCTCC 594
244 GGGATGTACCATGTGCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGAC 303
595 GGGATGTACCATGTGCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGAC 654
304 ATGATCATGCACACCCCGGGTGGCTGCTGCGTTCGGGAGAACAACTCTTCCGCTGC 363
655 ATGATCATGCACACCCCGGGTGGCTGCTGCGTTCGGGAGAACAACTCTTCCGCTGC 714
364 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGGTCGCCACCAACGACAATA 423
715 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGGTCGCCACCAACGACAATA 774
424 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTCTGCTATGATGCTGGGG 834
775 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTCTGCTATGATGCTGGGG 894
439 -----TCCAGCTGTTTACCATCTCGCTCGCGGCGAT 471
835 GACCTCTGCGGATCTGCTTCTCTCGTCTCCAGCTGTTTACCATCTCGCTCGCGGCGAT 894
472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCAACCGTATG 531
895 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCAACCGTATG 954
532 GCTTGGGATATGATGATGACTGCT 556
955 GCTTGGGATATGATGATGACTGCT 979

RESULT 14
AAT12706
ID AAT12706 standard; DNA; 633 BP.
XX
AC AAT12706;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC111A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.

304 ATGATCATGCACACCCCGGGTGGCTGCTGCGTTCGGGAGAACAACTCTTCCGCTGC 363
364 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGGTCGCCACCAACGACAATA 423
364 TGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGGTCGCCACCAACGACAATA 423
424 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTCTGCTATGATGCTGGGG 483
424 CGACGCCACGTGCTGATTTGCTTGGGGCGGCTGCTTCTGCTATGATGCTGGGG 483
439 -----TCCAGCTGTTTACCATCTCGCTCGCGGCGAT 471
484 GACCTCTGCGGATCTGCTTCTCTCCAGCTGTTTACCATCTCGCTCGCGGCGAT 543
472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCAACCGTATG 531
544 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCAACCGTATG 603
532 GCTTGGGATATGATGATGACTGCT 556
604 GCTTGGGATATGATGATGACTGCT 628

RESULT 13
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX
AC AAT12974;
XX
DT 25-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC166.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-BP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
PS WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
XX and E2 protein coding sequence constructs. These sequences are included
XX in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
XX The recombinant proteins can then be isolated using a method of the
XX invention. In the method, the envelope proteins are purified by
XX carrying out a disulphide bond cleavage, or a reduction step with a
XX disulphide bond cleavage agent, after lysis of recombinant host cells.
XX The constructs containing the purified HCV envelope proteins can be used
XX for vaccinating humans against HCV, for in vitro detection of HCV
XX antibodies in a sample, and in a serotyping assay for detecting one or
XX more serological types of HCV present in a biological sample. The
XX constructs can also be immobilised on a solid substrate and incorporated
XX into a reversed phase hybridisation assay for determining the presence or

PN WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of dl:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
SQ Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;

Query Match 82.2%; Score 461.4; DB 17; Length 633;
Best Local Similarity 86.9%; Pred. No. 2.2e-113;
Matches 550; Conservative 0; Mismatches 11; Indels 72; Gaps 1;

QY 1 ATGTGGTAAAGTCATCGATACCTTACATCGGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTGGTAAAGTCATCGATACCTTACATCGGGCTTCGCCGACCTCATGGGTACATT 60

QY 61 CCGCTCGTCGGCGCCCCCTTAGGGGGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTTAGGGGGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTT 120

QY 121 CTGGAGACGGGTGAATATGCAACAGGAATTTGCCGTGCTCTTCTATCTTC 180
DB 121 CTGGAGACGGGTGAATATGCAACAGGAATTTGCCGTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTGTGCTCTCTGACCGTTCCAGTTCCTGATGAGTGGCGCAAGTG 240
DB 181 CTCTTGGCTTGTGCTCTCTGACCGTTCCAGTTCCTGATGAGTGGCGCAAGTG 240

QY 241 TCCGGATGTACCATGTGACAGCAAGCTGTCCAACTCAAGCATTTGTGTATGAGGACGC 300
DB 241 TCCGGATGTACCATGTGACAGCAAGCTGTCCAACTCAAGCATTTGTGTATGAGGACGC 300

QY 301 GACATGATATGACACCCCGGGTGGCTCCCTTGGCTTCGGGAGAACAACTTCTCCCGC 360
DB 301 GACATGATATGACACCCCGGGTGGCTCCCTTGGCTTCGGGAGAACAACTTCTCCCGC 360

QY 361 TGCTGGTAGCGTCAACCCCGCTCGAGCTAGGAACGCGAGCGTCCCGACACACA 420
DB 361 TGCTGGTAGCGTCAACCCCGCTCGAGCTAGGAACGCGAGCGTCCCGACTACGACA 420

QY 421 ATACGACGCCACGTCGAT----- 438
DB 421 ATACGACGCCACGTCGATTTGCTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG 480
QY 439 -----TCCAGCTGTTCAACCATCTCGCTCGCGG 468
DB 481 GGGGATCTCTGCGGATCTGCTTCTCCGCTGCCAGCTGTTCCACCATCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 528
DB 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAAGCGGTCAACGT 600
QY 529 ATGCGTTGGGATATGATGATGAATGAACTGGTAATAG 561
DB 601 ATGCGTTGGGATATGATGATGAATGAACTGGTAATAG 633

RESULT 15
AAL48915
ID AAL48915 standard; DNA; 633 BP.
XX
AC AAL48915;
XX
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus clone HCC111A E1 protein coding sequence.
XX
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
PN WO200255548-A2.
XX
PD 18-JUL-2002.
XX
PF 11-JAN-2002; 2002WO-EP00219.
XX
PR 11-JAN-2001; 2001US-260699P.
PR 30-AUG-2001; 2001US-315768P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Maertens G, Bosman F, Buyse M;
XX
DR WPI; 2002-599657/64.
XX P-PSDB; AAO18662.
XX
PT New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX
PS Example 2; Page 163-164; 243pp; English.
XX
CC The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;

Query Match 82.2%; Score 461.4; DB 24; Length 633;
Best Local Similarity 86.9%; Pred. No. 2.2e-113;

Matches		550;	Conservative	0;	Mismatches	11;	Indels	72;	Gaps	1;
Qy	1	ATGTTGGGTAAGTATCATGATACCCCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT	60							
Db	1	ATGTTGGGTAAGTATCATGATACCCCTTACATCGGCTTCGCCGACCTCATGGGGTACATT	60							
Qy	61	CCGCTCGTGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120							
Db	61	CCGCTCGTGGCGCCCTAGGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGGTT	120							
Qy	121	CTGGAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTCTATCTTC	180							
Db	121	CTGGAGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTCTATCTTC	180							
Qy	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCGAACGCTG	240							
Db	181	CTCTTGGCTTTTACTGTCTCTGTGACCAATTCAGCTTCCGCTTATGAAGTGGCGAACGCTG	240							
Qy	241	TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGCG	300							
Db	241	TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGACGCG	300							
Qy	301	GACATGATCATGACACACCCCGGGTGGCTGCTGCGTTCCGGAGAACACTCTTCCCGC	360							
Db	301	GACATGATCATGACACACCCCGGGTGGCTGCTGCGTTCCGGAGAACACTCTTCCCGC	360							
Qy	361	TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGCGTCCCGCACGACA	420							
Db	361	TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGCGTCCCGCACGACA	420							
Qy	421	ATACGACGCCACGTCGAT-----	438							
Db	421	ATACGACGCCACGTCGATTTGCTGTTGGGGCGCTGCTTTCTGTTCGGCTATGTACGTG	480							
Qy	439	-----TCCGAGCTGTTCCACATCTCGCCTCGCGG	468							
Db	481	GGGGATCTCTGGGATCTGTCTCTGCTCTCCAGCTGTTCCACATCTCGCCTCGCGG	540							
Qy	469	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	528							
Db	541	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT	600							
Qy	529	ATGGCTTGGGATATGATGATGAACTGGTAATAG	561							
Db	601	ATGGCTTGGGATATGATGATGAACTGGTAATAG	633							

Search completed: December 19, 2003, 18:51:15
Job time : 158.093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1443.21 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-23
Perfect score: 561
Sequence: 1 ATGTTGGTAAGGTATCGA.....TGATGTAAGTGGTAATAG 561

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

```

1:  em_estba:*
2:  em_esthum:*
3:  em_estin:*
4:  em_estnu:*
5:  em_estov:*
6:  em_estpl:*
7:  em_estro:*
8:  em_htc:*
9:  gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vxl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	62.6	11.2	488	9 AV755731	AV755731 AV755731
C 2	50.6	9.0	492	9 AV758366	AV758366 AV758366
C 3	41.6	7.4	502	12, BI879124	BI879124 fm04608.y
C 4	40.6	7.2	275	9 AV835132	AV835132 AV835132

5	40.6	7.2	402	9	AV392783	AV392783 AV392783
6	40.6	7.2	551	9	AV392165	AV392165 AV392165
7	40.6	7.2	552	12	BI996341	BI996341 1031037A0
8	40.6	7.2	584	12	BI727879	BI727879 1031095C1
9	40.4	7.2	1201	13	BI356664	BI356664 BX356664
10	40.2	7.2	1162	12	BM918259	BM918259 AGENCOURT
C 11	40	7.1	1201	9	AL513886	AL513886 AL513886
C 12	39	7.0	359	12	BJ252869	BJ252869 BJ252869
13	39	7.0	375	12	BJ246716	BJ246716 BJ246716
14	39	7.0	840	29	CC335916	CC335916 OGUAJ60TV
15	39	7.0	873	14	CD446071	CD446071 EL0170207
16	38.6	6.9	925	29	CNS0091P	AL053013 Drosophila
17	38.4	6.8	636	12	BI960110	BI960110 HVSME002
18	38.4	6.8	702	14	CD432549	CD432549 ETH1_30_D
19	38.4	6.8	970	29	CNS010C9	AL098787 Drosophila
C 20	38.4	6.8	987	29	CNS015VX	AL105975 Drosophila
C 21	38.2	6.8	533	6	AU192776	AU192776 Porphyra
C 22	38.2	6.8	538	6	AU193705	AU193705 Porphyra
C 23	38.2	6.8	544	6	AU190971	AU190971 Porphyra
C 24	38.2	6.8	1270	12	BG968359	BG968359 602834809
C 25	38.2	6.8	1270	12	CG966525	CG966525 NL38_G12
C 26	38	6.8	1201	13	BX381961	BX381961 BX381961
C 27	38	6.8	1201	13	BX381961	BX381961 BX381961
28	37.8	6.7	435	14	C72860	C72860 C72860 Rice
29	37.8	6.7	533	29	CC010084	CC010084 FUGJ092TB
30	37.8	6.7	659	29	CC405164	CC405164 PUHSP17TB
31	37.8	6.7	826	29	BZ736582	BZ736582 OGER42TC
C 32	37.8	6.7	895	29	CC359028	CC359028 PUHFD18TD
C 33	37.8	6.7	925	29	CC359026	CC359026 PUHFD18TB
C 34	37.8	6.7	940	29	CC010085	CC010085 FUGJ092TD
C 35	37.8	6.7	951	29	CC405167	CC405167 PUHSP17TD
36	37.6	6.7	431	9	AV639153	AV639153 AV639153
37	37.4	6.7	360	9	AJ473805	AJ473805 AJ473805
C 38	37.4	6.7	637	13	BQ293470	BQ293470 1091016FO
C 39	37.4	6.7	641	13	BQ172543	BQ172543 1091024A1
C 40	37.4	6.7	650	14	CA828039	CA828039 1114022H0
41	37.4	6.7	834	29	BZ641450	BZ641450 OGCCS41TC
C 42	37.4	6.7	841	29	BZ641457	BZ641457 OGCCS41TM
C 43	37.4	6.7	856	29	BZ578381	BZ578381 msh2_5817
44	37.4	6.7	872	29	BZ555011	BZ555011 pacsl-60
45	37.2	6.6	563	10	BE490055	BE490055 WHE0364_A

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

```

source
1. .488
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="BMFAK803"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      116 a   134 c   137 g   97 t       4 others
ORIGIN

Query Match      11.2%; Score 62.6; DB 9; Length 488;
Best Local Similarity 77.9%; Pred. No. 1e-05;
Matches 88; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 445 CTGTTCCACCATCTGGCCTCGCCGGCATGACAGCGTGCAGGACTGCAATTCGTCACATCTAT 504
DB 403 CAGCTGATCATCTGGCCTCAGCACCATGAGTTGTGTCATGATGCAACTGCTCCATCTAT 344

QY 505 CCCGGCCACATACCGGTCACCGCTATG-GCTTGGGATGATGATGATGATGATGATGATGAT 556
DB 343 CCTGGCCCATCACTGACACCGCATGAGCATGGACATGATGATGATGATGATGATGATGATGAT 291

RESULT 2
AV758366/c      492 bp      mrna      linear      EST 19-OCT-2000
LOCUS      AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION      AV758366
ACCESSION      AV758366
VERSION      AV758366.1 GI:10916214
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 492)
AUTHORS      Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
               Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,
               L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
               Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
TITLE      Homo sapiens cDNA BM clones
JOURNAL      Unpublished
COMMENT      Contact: Zequang Han
               Chinese National Human Genome Center at Shanghai
               351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai-
               201203, P. R. China
               Tel: 86-21-50801919 (ex. 45)
               Fax: 86-21-50801922
               Email: hanzg@chgc.sh.cn
               This clone is available at CHGC in Shanghai.
FEATURES
source
1. .492
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      124 a   128 c   125 g   112 t       3 others
ORIGIN

Query Match      9.0%; Score 50.6; DB 9; Length 492;
Best Local Similarity 72.5%; Pred. No. 0.013;
Matches 79; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 449 TCACCATCTGCCGCGCCGATGACAGCGTGCAGGACTGCAATTCGTCACATCTATCCCG 508

```



```

Db      364 CTACCGCGGAGGTGGCTAAGTAGCTTACAGGTTGACGCTCAGTCGGCAGAACCGAA 305
QY      480 GCAGGACTGCAA 491
Db      304 GTAGCGGTGCAA 293

RESULT 4
AV835132
LOCUS   AV835132
DEFINITION AV835132 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
spontaneum top three leaves adult, heading stage Hordeum vulgare
subsp. spontaneum cDNA clone bah24018, mRNA sequence.
ACCESSION AV835132
VERSION   AV835132.1 GI:14527221
KEYWORDS  EST.
SOURCE    Hordeum vulgare subsp. spontaneum
ORGANISM  Hordeum vulgare subsp. spontaneum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
          ; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 275)
AUTHORS   Sato, K.
TITLE     Barley EST sequencing project in NIG and Okayama Univ
JOURNAL   Unpublished
COMMENT   Contact: Kazuhiro Sato
          Research Institute for Bioresources
          Okayama University, Barley Germplasm Center
          Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
          Email: kazeato@rib.okayama-u.ac.jp/
          URL: http://www.rib.okayama-u.ac.jp/barley/
          Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct
          submission;
          database: http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES             source
ORIGIN               1..275
                    /organism="Hordeum vulgare subsp. spontaneum"
                    /mol_type="mRNA"
                    /cultivar="H602"
                    /db_xref="taxon:77009"
                    /clone="bah24018"
                    /issue_type="top three leaves"
                    /dev_stage="adult, heading stage"
                    /clone_lib="K. Sato unpublished cDNA library: Hordeum
                    vulgare subsp. spontaneum top three leaves adult, heading
                    stage"
BASE COUNT          30 a 123 c 67 g 46 t 9 others
ORIGIN

Query Match          7.2%; Score 40.6; DB 9; Length 275;
Best Local Similarity 48.2%; Pred. No. 4.3;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY      306 GATCATGCACACCCCGGTGGTGGCCCTGCGTTCGGGAGAACAACTTCCCGCTGCTG 365
Db      1  GGTCTCTCGACGGAACCCGCCCTCTGCTCCGTTCGTCTCCCTGGCTGCGCACGCCCG 60

QY      366 GGTAGCGCTCACCCACGCTCGCAGCTAGGACGCGCAGCGTCCCCACCAACAATACG 425
Db      61  CGGCCAGCGGCACTCACCTNCTCTGCTCGTTGCGCCCTCCCTCTCGCGCCGCCCTCG 120

QY      426 ACGCCACGTCGATTCCTCCAGCTGTTCACCATCTCGCTCGCGGCATGACGCGTGCAGGA 485
Db      121 ACGGCTGGCGTAGCTGGGTGAGTTACATCTCTGCTCTCCGCGAGCGGACGCTGCGCGA 180

QY      486 CTGCAATTGCTCAATCTATCCCGGCCACATACACGGGTCAACCGTATG 531
Db      181 ACGGACTCGCCTCCCTCTCCGCCCCAGCGCTCGCACCGGAACG 226

RESULT 5
AV392783
LOCUS   AV392783
DEFINITION AV392783 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM096g04_r 5', mRNA sequence.
ACCESSION AV392783
VERSION   AV392783.1 GI:6546999
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
          Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
          Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 402)
AUTHORS   Aamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE     A large scale structural analysis of cDNAs in a unicellular green
          alga, Chlamydomonas reinhardtii. I. Generation of 3433
          non-redundant expressed sequence tags
JOURNAL   DNA Res. 6 (6), 369-373 (1999)
MEDLINE   20152988
PUBMED    10691129
COMMENT   Contact: Yasukazu Nakamura
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES             source
ORIGIN               66 a 126 c 146 g 64 t
                    /organism="Chlamydomonas reinhardtii"
                    /mol_type="mRNA"
                    /strain="C9"
                    /db_xref="taxon:3055"
                    /clone="CM096g04_r"
                    /dev_stage="photoautotrophic growth"
                    /clone_lib="Chlamydomonas reinhardtii C9"
                    /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                    XhoI"
BASE COUNT          66 a 126 c 146 g 64 t
ORIGIN

Query Match          7.2%; Score 40.6; DB 9; Length 402;
Best Local Similarity 45.3%; Pred. No. 4.7;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY      42  CGACCTCTGGGTACATTCCTCGCTCGCGCGCCCTAGGGGGGCGCTGCCAGGGCCCT 101
Db      53  CGAGCTCATCTCTGTCATTTGTGCGCGCACTGCAACATGAAGGACGCTGTGACGGACCT 112

QY      102 GCGCATGCGCTCCGGGTTCTGAGGAGCGCGTGAACATATGCAACAGGAATTGCCCGG 161
Db      113 GCGCGCGCGCGCGAGTGGGAGGGCGGCTACGCGCACGATCCGTGAGCTTGGGCGC 172

QY      162 TTGCTCTTTCTCTATCTTCTCTTGGCTTTGCTGTCTGTGACCGTTCACGCTTCCGC 221
Db      173 CGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGACCTCAAGGCCAGAACCCAG 232

QY      222 TTATGATGCGCAACGTCGCGGATGACCATGTACACGAAAGACTGCTTCAACTCAAG 281
Db      233 CTTTCGCGCTCGCTGCTGGGCGCACCTCGTGGCGCGGCGCACCGCGCTGCTGTGAT 292

QY      282 CATTGTGTATGAGCGAGCGACATGATCATGACACCCCGGCTGCGTGCCTCGCTCG 341
Db      293 CTTGATGACCAAGAGAGTTCGGGCGCGCATCTACGGCGGCGGTGCCATGCGCGG 352

QY      342 GGAGAACAACTTCTCCCGCTGCTGGGT 368
Db      353 CAGAGAGCAAGGCGAGCTACATGAT 379

RESULT 6
AV392165
LOCUS   AV392165
DEFINITION AV392165 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
cDNA clone CM083e05_r 5', mRNA sequence.
ACCESSION AV392165
VERSION   AV392165.1 GI:6546381

```

RESULT	7
B1996341	
LOCUS	
DEFINITION	552 bp mRNA linear EST 25-OCT-2001 1011037A07.Y2 C. reinhardtii CC-1690, Stress II (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION	B1996341
VERSION	B1996341.1 GI:16431115
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii
ORGANISM	Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.

QY	394	AGGAAGCGCAGCTGCCACGACGATACGACCGCGATGATCCCGCTGTTCCAC	453						
Db	975	STTBSTTTTSSBTTSSBTTTSSBTTTSSBTTSSBTTSSBTTSSBTTSSBTTSS	1034						
QY	454	ATCTCGCTCCGCGCATGAGCGGTGAGGACTGCAATGCTCAATCTATCCCGCCAC	513						
Db	1035	SNKSSSSSSSSSSSSSSSSSSSSBTTTSSBTTSSBTTSSBTTSSBTTSSBTTSS	1094						
QY	514	ATAACGGGTCCAGGTATGCC	533						
Db	1095	SBTTSTTTTSSSATTSSB	1114						
RESULT 10									
BM918259		1162 bp	linear	EST 12-MAR-2002					
LOCUS									
DEFINITION	AGENCOURT 6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649								
ACCESSION	BM918259								
VERSION	BM918259.1	GI:19368638							
KEYWORDS	EST.								
SOURCE	Homo sapiens (human)								
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
REFERENCE	1 (bases 1 to 1162)								
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/								
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)								
JOURNAL	Unpublished								
COMMENT	Contact: Robert Strauberg, Ph.D. Email: csapbes-remail.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2016 row: n column: 18 High quality sequence stop: 567.								
FEATURES									
source									
1..1162									
/organism="Homo sapiens"									
/mol_type="mRNA"									
/db_xref="taxon:9606"									
/clone="IMAGE:5485649"									
/tissue_type="natural killer cells, cell line"									
/lab_host="DH10B (phage-resistant)"									
/clone_lib="NIH MGC 106"									
/note="Organ: blood; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."									
BASE COUNT	224 a	499 c	240 g	198 t	1 others				
ORIGIN									
Query Match	7.2%	Score 40.2;	DB 12;	Length 1162;					
Best Local Similarity	54.4%;	Pred. No. 7.9;							
Matches	81;	Conservative	0;	Mismatches	68;	Indels	0;	Gaps	0;
QY	313	CACACCCCGGGTGGCGCTCGCTTCGGGAGAACAACTTTCCCGCTGCTGGGTAGCG	372						
Db	715	CCCCCCCCCGGGTCCCTGCCCCACACCCCGGACCCCAAAACCCCGCGGACCGCTCC	774						
QY	373	CTCACCCCGCAGCTCGGACGTAGGACGCGTCCCAACACGACAAATACGACGAC	432						
Db	775	CCCACTCGACCGCACCCCGCCCATATCCGCCCCCCCCCTACCGGATACCACTACCCAC	834						
QY	835	GCGTATCCCGCGCTGACACACCCCGCC	863						

```

QY 332 CTTGCGTTCGGAGAACAACTCTTCGCGTCTGGGTAGCGTCAACCCACGCGTCCGAG 391
Db 832 SGCCKGKTKTKTTTGTGCGTTGAAGASABAGTWTGAGGGGGGGCCCYCSCMCCCCCYYB 773
QY 392 CTAGGAACAGCGAGCGTCCCCACACGACA 420
Db 772 BBCCCMCHCTKCSCKMCCRCGACTYCCCA 744

RESULT 12
BJ252669/c
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
BJ252669
VERSION
GI:20061830
KEYWORDS
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 359)
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..359
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_f"
/clone 115 c 107 g 67 t

BASE COUNT 70 a 115 c 107 g 67 t
ORIGIN

Query Match 7.0%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CTTACATGCGGCTTCGCGACCTCTGGGGTACATTCCGCTCGCGCGCCCCCTAGG 83
Db 297 CTTCAAGTGAACAGCGCGTCTGGAAGCGCTCAGGGCGTGCAGCGCGTCCGCGCG 238
QY 84 GGGCGCTGCAGGCGCTGCGCATGCGCTCGGGTTCTGGAGAGCGCGTGAACATATG 142
Db 237 GGACGCGCGCAGCGCCCTGGCGCAGGACGTCGACGCTGCGCGTGCACGTGCCCAAGG 179

RESULT 13
BJ246716
LOCUS
DEFINITION
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
BJ246716
VERSION
GI:20058228
KEYWORDS
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 375)
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_f"
/clone 115 c 107 g 67 t

BASE COUNT 70 a 115 c 107 g 67 t
ORIGIN

Query Match 7.0%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CTTACATGCGGCTTCGCGACCTCTGGGGTACATTCCGCTCGCGCGCCCCCTAGG 83
Db 297 CTTCAAGTGAACAGCGCGTCTGGAAGCGCTCAGGGCGTGCAGCGCGTCCGCGCG 238
QY 84 GGGCGCTGCAGGCGCTGCGCATGCGCTCGGGTTCTGGAGAGCGCGTGAACATATG 142
Db 237 GGACGCGCGCAGCGCCCTGGCGCAGGACGTCGACGCTGCGCGTGCACGTGCCCAAGG 179

RESULT 14
CC335916
LOCUS
DEFINITION
genomic survey sequence.
ACCESSION
CC335916
VERSION
GI:30805329
KEYWORDS
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0393124"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site:1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

```

Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Feekes' scale 10.5.1"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_f"
/clone 107 c 110 g 77 t

BASE COUNT 81 a 107 c 110 g 77 t
ORIGIN

Query Match 7.0%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CTTACATGCGGCTTCGCGACCTCTGGGGTACATTCCGCTCGCGCGCCCCCTAGG 83
Db 36 CTTCAAGTGAACAGCGCGTCTGGAAGCGCTCAGGGCGTGCAGCGCGTCCGCGCG 95
QY 84 GGGCGCTGCAGGCGCTGCGCATGCGCTCGGGTTCTGGAGAGCGCGTGAACATATG 142
Db 96 GGACGCGCGCAGCGCCCTGGCGCAGGACGTCGACGCTGCGCGTGCACGTGCCCAAGG 154

RESULT 14
CC335916
LOCUS
DEFINITION
genomic survey sequence.
ACCESSION
CC335916
VERSION
GI:30805329
KEYWORDS
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0393124"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site:1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

Mon Dec 22 13:28:41 2003

Search completed: December 20, 2003, 06:54:49
Job time : 1444.21 secs

```
BASE COUNT      103 a      298 c      291 g      148 t
ORIGIN
Query Match      7.0%; Score 39; DB 29; Length 840;
Best Local Similarity 53.6%; Pred. No. 15;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 39 CGCCGACCTGTGGGGTACATTCCGCTCGTCGGCGCCCTAGGGGGGCTGCCAGGGC 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CCCGGCGTCGACGGGCCACTCCGCTTCACTGCGCGGCTCCGGCGGCTCGCGAACGC 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 99 CTTGGCGCATGGGTCGGGTTCTGGAGGACGGCGTGAATATGCAACAGGGAATTGCC 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 TGTGGCGCTGCTCAAGCTGCTCTGGTGGCGGCGCGACCCGGTAATCCCGACTCCGC 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 159 CGGTTGCTCTTTCTATCTTCCTTCTTGGCT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 CGGTCGCTTGCTGCCGACGTCATCTTGCT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
CD446071      873 bp      mRNA      linear      EST 03-JUN-2003
DEFINITION    EL01T0207B11.b EndospERM_4 Zea mays cDNA, mRNA sequence.
ACCESSION     CD446071
VERSION       CD446071.1 GI:31361714
KEYWORDS      EST.
SOURCE        Zea mays
ORGANISM      Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 873)
AUTHORS       Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
              Messing, J.
TITLE         Sequencing of the maize endospERM ESTs
JOURNAL       Unpublished
COMMENT       Contact: Lai, Jinsheng
              Dr. Joachim Messing's lab
              Wakeman Institute, Rutgers University
              190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
              Tel: 732-445-3801
              Fax: 732-445-5735
              Email: jlai@wakeman.rutgers.edu
              Seq primer: T3.
              Location/Qualifiers
                1..873
                  /organism="Zea mays"
                  /mol_type="mRNA"
                  /cultiivar="W22"
                  /db_xref="taxon:4577"
                  /issue_type="EndospERM of 7-23DAP"
                  /clone_lib="EndospERM 4"
                  /notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                  XhoI"

BASE COUNT      109 a      300 c      294 g      170 t
ORIGIN
Query Match      7.0%; Score 39; DB 14; Length 873;
Best Local Similarity 53.6%; Pred. No. 15;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 39 CGCCGACCTGTGGGGTACATTCCGCTCGTCGGCGCCCTAGGGGGGCTGCCAGGGC 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 583 CCCGGCGTCGACGGGCCACTCCGCTTCACTGCGCGGCTCCGGCGGCTCGCGAACGC 642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 99 CTTGGCGCATGGGTCGGGTTCTGGAGGACGGCGTGAATATGCAACAGGGAATTGCC 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 643 TGTGGCGCTGCTCAAGCTGCTCTGGTGGCGGCGCGACCCGGTAATCCCGACTCCGC 702
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 159 CGGTTGCTCTTTCTATCTTCCTTCTTGGCT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 703 CGGTCGCTTGCTGCCGACGTCATCTTGCT 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 40.4591 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-23

Perfect score: 561

Sequence: 1 ATGTTGGTAAAGTCATCGA.....TGATGATGAACCTGGTAATAG 561

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	561	100.0	561	3	US-08-612-973-23
2	561	100.0	561	3	US-08-927-597-23
3	556	99.1	606	3	US-08-612-973-25
4	556	99.1	606	3	US-08-927-597-25
5	556	99.1	636	3	US-08-612-973-27
6	556	99.1	636	3	US-08-927-597-27
7	556	99.1	723	3	US-08-612-973-21
8	556	99.1	723	3	US-08-927-597-21
9	474	84.5	795	3	US-08-612-973-5
10	474	84.5	795	3	US-08-927-597-5
11	471	84.0	2082	3	US-08-612-973-47
12	471	84.0	2082	3	US-08-927-597-47
13	471	84.0	2433	3	US-08-612-973-49
14	471	84.0	2433	3	US-08-927-597-49
15	461.4	82.2	633	3	US-08-612-973-7
16	461.4	82.2	633	3	US-08-927-597-7
17	454	80.9	636	3	US-08-612-973-13
18	454	80.9	636	3	US-08-927-597-13
19	415.4	74.0	1539	2	US-08-470-426B-17
20	415.4	74.0	1863	2	US-08-470-426B-14
21	413.4	73.7	1037	1	US-08-462-195-1
22	413.4	73.7	1037	1	US-08-636-883-1
23	413.4	73.7	1037	3	US-09-127-829-1
24	410.6	73.2	932	1	US-08-081-072-15
25	410.6	73.2	932	1	US-08-449-093A-15
26	410.6	73.2	9595	3	US-09-014-416-4
27	410.6	73.2	9599	3	US-09-014-416-6

Sequence 21, Appl
Sequence 96, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl

US-08-191-160-21
US-08-150-204E-96
US-08-324-977-9
US-08-384-616-9
US-08-904-686A-9
US-09-315-850-9
US-08-324-977-3
US-08-384-616-3
US-08-904-686A-3
US-09-315-850-3
US-08-324-977-11
US-08-384-616-11
US-08-904-686A-11
US-09-315-850-11
US-08-324-977-13
US-08-384-616-13
US-08-904-686A-13
US-09-315-850-13

ALIGNMENTS

RESULT 1
US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555


```

RESULT 3
US-08-612-973-25
; Sequence 25, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

```

Query Match	99.1%; Score 556; DB 3; Length 606;
Best Local Similarity	100.0%; Pred. No. 2.5e-140;
Matches 556; Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT 60
DB	1 ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT 60
QY	61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGGTT 120
DB	61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGGTT 120
QY	121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAAATTTGCCCGGTTGCTCTTCTCTATATCTTC 180
DB	121 CTGGAGGACGGCGTGAACCTATGCAACAGGGNAATTTGCCCGGTTGCTCTTCTCTATCTTC 180
QY	181 CTCCTTGGCTTTGCTGTCCTGTCTGACCGGTTCCAGCGTTCCGGCTTATGAAGTCCGCAACGTG 240
DB	181 CTCCTTGGCTTTGCTGTCCTGTCTGACCGGTTCCAGCTTCCGGCTTATGAAGTCCGCAACGTG 240

RESULT 4
US-08-927-597-25
Sequence 25, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:

Db 541 ATGATGATGAAGTGGT 556

RESULT 6
US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630
US-08-927-597-27

Query Match 99.1%; Score 556; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.5e-140;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTGCATCGATACCCCTTACATGGCGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAGTGCATCGATACCCCTTACATGGCGGCTTCGCCGACCTCGTGGGGTACATT 60

QY 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCGTCCGGGTT 120
DB 61 CGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCGTCCGGGTT 120

QY 121 CTGGAGAGCGCGTGAAGTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGAGCGCGTGAAGTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGTCTGTCTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACAGTG 240
DB 181 CTCTTGGCTTTGCTGTCTGTCTGACCGCTTCCAGCTTCCGCTTATGAAGTGCACAGTG 240

QY 241 TCGGGATGTACCATGTACGACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300
DB 241 TCGGGATGTACCATGTACGACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300

QY 301 GACATGATCATGCACACCCCGGGTGCCTCGCTTCCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGCCTCGCTTCCGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGTAGCGCTACCCCGGCTCGCAGCTAGGAAGCCAGCGTCCCGACGACA 420
DB 361 TGCTGGTAGCGCTACCCCGGCTCGCAGCTAGGAAGCCAGCGTCCCGACGACA 420

QY 421 ATACGACGCCAGCTCGATTCCCGAGCTGTTCAACATCTCGCTCGCGGCATGAGCGTG 480
DB 421 ATACGACGCCAGCTCGATTCCCGAGCTGTTCAACATCTCGCTCGCGGCATGAGCGTG 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGAT 540

QY 541 ATGATGATGAAGTGGT 556
DB 541 ATGATGATGAAGTGGT 556

RESULT 7
US-08-612-973-21
; Sequence 21, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO

Mon Dec 22 13:28:39 2003

```

; ANTI-SENSE: NO
; FEATURE: CDS
; NAME/KEY: 1..720
; LOCATION: 1..720
; FEATURE: mat_peptide
; NAME/KEY: 1..717
; LOCATION: 1..717
; US-08-612-973-21

Query Match 99.1%; Score 556; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.6e-140;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCATCGATACCCCTACATCGGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCCCTACATCGGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCCGGGTT 120
DB 61 CCGCTCGTTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCCGGGTT 120
QY 121 CTGAGAGCGCGTGAATATGCAAGGAAATTCGCGGTTGCTTTCTATCTTC 180
DB 121 CTGAGAGCGCGTGAATATGCAAGGAAATTCGCGGTTGCTTTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
DB 181 CTCTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGAAAGATGCTGCCAATCAAGCATTTGTATGAGGAGCG 300
DB 241 TCCGGGATGTACCATGTCAAGAAAGATGCTGCCAATCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
DB 301 GACATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAGCGCAGCGTCCCAACACACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAGCGCAGCGTCCCAACACACA 420
QY 421 ATACAGCGCGAGTCCAGTCTCCAGCTGTTCCACATCTCCGCTCCGCGCATGACAGCGTG 480
DB 421 ATACAGCGCGAGTCCAGTCTCCAGCTGTTCCACATCTCCGCTCCGCGCATGACAGCGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACGTGT 556
DB 541 ATGATGATGAACGTGT 556
```

```

RESULT 8
US-08-927-597-21
; Sequence 21, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..717
; US-08-927-597-21

Query Match 99.1%; Score 556; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.6e-140;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCATCGATACCCCTACATCGGCTTCGCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTAAAGTCATCGATACCCCTACATCGGCTTCGCGACCTCGTGGGTACATT 60
QY 61 CCGCTCGTTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCCGGGTT 120
DB 61 CCGCTCGTTCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCCGGGTT 120
QY 121 CTGAGAGCGCGTGAATATGCAAGGAAATTCGCGGTTGCTTTCTATCTTC 180
DB 121 CTGAGAGCGCGTGAATATGCAAGGAAATTCGCGGTTGCTTTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
DB 181 CTCTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGAAAGATGCTGCCAATCAAGCATTTGTATGAGGAGCG 300
DB 241 TCCGGGATGTACCATGTCAAGAAAGATGCTGCCAATCAAGCATTTGTATGAGGAGCG 300
QY 301 GACATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
DB 301 GACATGATCATGCAACACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAGCGCAGCGTCCCAACACACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAGCGCAGCGTCCCAACACACA 420
QY 421 ATACAGCGCGAGTCCAGTCTCCAGCTGTTCCACATCTCCGCTCCGCGCATGACAGCGTG 480
DB 421 ATACAGCGCGAGTCCAGTCTCCAGCTGTTCCACATCTCCGCTCCGCGCATGACAGCGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540
```

Db	481	CAGACATGCAATGCTCAATCATATCCGCGCCACATACGGGTACCGGTATGCTGGGAT	556
Qy	541	ATGATGATGAACTGGT	556
Db	541	ATGATGATGAACTGGT	556
RESULT 9			
US-08-612-973-5			
; Sequence 5, Application US/08612973			
; Patent No. 6150134			
; GENERAL INFORMATION:			
; APPLICANT: MAERTENS, GEERT			
; APPLICANT: BOSMAN, FONS			
; APPLICANT: DE MARTYNOFF, GUY			
; APPLICANT: BUYSE, MARIE-ANGE			
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE			
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE			
; NUMBER OF SEQUENCES: 111			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: NIXON & VANDERHYE P. C.			
; STREET: 1100 NORTH GLEBE ROAD			
; CITY: ARLINGTON			
; STATE: VIRGINIA			
; COUNTRY: U.S.A.			
; ZIP: 22201-4714			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/612,973			
; FILING DATE: 11-MAR-1996			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: BYRNE, THOMAS E.			
; REGISTRATION NUMBER: 32,205			
; REFERENCE/DOCKET NUMBER: 1487-10			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (703) 816-4000			
; TELEFAX: (703) 816-4100			
; INFORMATION FOR SEQ ID NO: 5:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 795 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: CDNA			
; HYPOTHETICAL: NO			
; ANTI-SENSE: NO			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 1..792			
; FEATURE:			
; NAME/KEY: mat peptide			
; LOCATION: 1..789			
US-08-612-973-5			
Query Match 84.5%; Score 474; DB 3; Length 795;			
Best Local Similarity 88.5%; Pred. No. 2.8e-118;			
Matches 556; Conservative 0; Mismatches 0; Indels 72; Gaps 1			
Qy	1	ATGTTGGTAAAGTCATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGTAAAGTCATCGATACCCCTTACATCGGGCTTCGCCGACCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTCGCGCGCCCTAGGGGGCGCTGCAGGGCCCTCGCGCATGGCGTCCGGGTT	120
Db	61	CCGCTCGTCGCGCGCCCTAGGGGGCGCTGCAGGGCCCTCGCGCATGGCGTCCGGGTT	120
Qy	121	CTGAGGACCGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTCTATCTTC	180

TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
-08-08-597-5

Query Match	84.5%;	Score 474;	DB 3;	Length 795;
Best Local Similarity	88.5%;	Prid. No. 2.8e-118;		
Matches 556;	Conservative	0;	Mismatches 0;	Indels 72;
Gaps	1;			
1	ATGTTGGGTAAAGGTCATCGATAACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60		
y				
b	1	ATGTTGGGTAAAGTTCATCGATAACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60	
61	CGGCTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120		
y				
b	61	CGGCTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120	
121	CTGGAGGACGGCTGAACTATGCAACAGGGAAATTTGCCGGTGTCTTTCTCTATCTTC	180		
y				
b	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAAATTTGCCGGTGTCTTTCTCTATCTTC	180	
181	CTCTTGGCTTTGTCGTCGTGTGACGGTTTCAGCTTCGGCTTATGAAGTGGCAACGTG	240		
y				
b	181	CTCTTGGCTTTGTCGTCGTGTGACGGTTTCAGCTTCGGCTTATGAAGTGGCAACGTG	240	
241	TCCGGGATGTACCATGTACAGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG	300		
y				
b	241	TCCGGGATGTACCATGTACAGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG	300	
301	GACATGATCATGACACACCCCGGGTGCGTGCCCTCGGTTCCGGAGAACAACTCTTCCCGC	360		
y				

301	GACATGATCATGACACACCCCGGGTGGTGGCTTCTGGTTCGGGAGAACAACTCTCTCCCGG	388
361	TGCTGGGTAGCGCTCACCCCAACGGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACA	420
361	TGCTGGGTAGCGCTCACCCCAACGGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACA	420
421	ATAGACGCCACGCTCGAT-----	438
421	ATAGACGCCACGCTCGATTTGCTCGTTGGGGGGCTGCTTCTGTTCCGCTATGTACGTG	480
439	-----TCCACGCTGTACCACTCTCGCTCCCGG	468
481	GGGACCTCTCGGATCTGTCTTCTCGTCTCCCGCTGTTACCACTCTCGCTCCCGG	540
489	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAAACGGGTCAACCT	528
541	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAAACGGGTCAACCT	600
529	ATGGCTTCGGGATATGATGATCAACTGGT	556
601	ATGCTTTGGGATATGATGATGAACCTGT	628

RESULT 11
US-08-612-9773-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT

APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2076

19-08-612-973-47

	Query Match	84.0%;	Score 471;	DB 3;	Length 2082;
	Best Local Similarity	88.5%;	Pred. No. 2.4e-117;		
	Matches 553;	Conservative	0;	Mismatches	72; Gaps 1;
QY	4	TTGGGTAAGGTCAATCGATACCCCTTACATCGGCGCTTCGGCGACCTCGTGGGTACATTCG	63		
Db	4	TTGGGTAAGGTCAATCGATACCCCTTACATCGGCGCTTCGGCGACCTCGTGGGTACATTCG	63		
QY	64	CTCGTCGGCGCCCCCTTAGGGGCGCTGGCAGGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123		
Db	64	CTCGTCGGCGCCCCCTTAGGGGCGCTGGCAGGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123		
QY	124	GAGGACGGCGTGAACTATGCAA CAGGGAAATTTGCCCGGTTCCTCTTCTATCTTCCTC	183		
Db	124	GAGGACGGCGTGAACTATGCAA CAGGGAAATTTGCCCGGTTCCTCTTCTATCTTCCTC	183		
QY	184	TTGGGCTTTGCTGTCTGTCTGACGTTTCACGCTTCGGCTTATGAAGTGGCGAACGTGCC	243		
Db	184	TTGGGCTTTGCTGTCTGTCTGACGTTTCACGCTTCGGCTTATGAAGTGGCGAACGTGCC	243		
QY	244	GGGATGTACCATGTCA CGAACGACTGCTCCAA CTCGAAGCATTTGTATGAGCGACGGAC	303		
Db	244	GGGATGTACCATGTCA CGAACGACTGCTCCAA CTCGAAGCATTTGTGTATGAGCGACGGAC	303		
QY	304	ATGATCATGCA CACCCCGGTGTGTCCTCGGTTTCGGGAGAAACAACCTCTTCCCGCTGC	363		
Db	304	ATGATCATGCA CACCCCGGTGTGTCCTCGGTTTCGGGAGAAACAACCTCTTCCCGCTGC	363		

QY 364 TGGGTAGCGCTACCCCGCAGCTCGCAGCTAGGAAAGCCAGCGTCCCGCAGCAATA 423
Db 364 TGGGTAGCGCTACCCCGCAGCTCGCAGCTAGGAAAGCCAGCGTCCCGCAGCAATA 423
QY 424 CGACGCCACGCTCGAT----- 438
Db 424 CGACGCCACGCTCGATTTGCTGTTGGGGGGCTGCTTCTGTTCCGCTATGATGAGTGGGG 483
QY 439 -----TCCAGCTGTTTCAACCATCTCGCCTCGCCGGGCAT 471
Db 484 GACCTCTGCGGATCTGTTCTTCCTCGTCTCCAGCTGTTTCAACCATCTCGCCTCGCCGGGCAT 543
QY 472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 531
Db 544 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 603
QY 532 GCTTGGGATATGATGAACTGGT 556
Db 604 GCTTGGGATATGATGAACTGGT 628

RESULT 12

US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-927-597-47

Query Match 84.0%; Score 471; DB 3; Length 2082;
Best Local Similarity 88.5%; Pred. No. 2.4e-117; Indels 72; Gaps 1;
Matches 553; Conservative 0; Mismatches 0;
QY 4 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCCG 63
Db 4 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCCG 63
QY 64 CTCGTCCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATCGCGGTCTG 123
Db 64 CTCGTCCGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGATCGCGGTCTG 123
QY 124 GAGGACGGCGTGAACATATCAACAGGGAATTTGCCCGGTTGCTTCTTCTATCTTCCCTC 183
Db 124 GAGGACGGCGTGAACATATCAACAGGGAATTTGCCCGGTTGCTTCTTCTATCTTCCCTC 183
QY 184 TTGGCTTTGCTGTCTGTCTGACCCGTTCCAGCTTTCGAGTTCGCGCAACGTGTCC 243
Db 184 TTGGCTTTGCTGTCTGTCTGACCCGTTCCAGCTTTCGAGTTCGCGCAACGTGTCC 243
QY 244 GGGATGTACATGTTCACGAACGACTGCTCCAATCAAGCATTTGTGTATGAGGCGGAC 303
Db 244 GGGATGTACATGTTCACGAACGACTGCTCCAATCAAGCATTTGTGTATGAGGCGGAC 303
QY 304 ATGATCATGCACACCCCGGGTGGCGCTCGGTTCCGGGAGAACACTTTCCTCCCGTGC 363
Db 304 ATGATCATGCACACCCCGGGTGGCGCTCGGTTCCGGGAGAACACTTTCCTCCCGTGC 363
QY 364 TGGGTAGCGCTCACCCCGCAGCTCGCAGCTAGGAAGCCAGCGTCCCGCAGCAATA 423
Db 364 TGGGTAGCGCTCACCCCGCAGCTCGCAGCTAGGAAGCCAGCGTCCCGCAGCAATA 423
QY 424 CGACGCCACGTCGAT----- 438
Db 424 CGACGCCACGTCGATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGATGAGTGGGG 483
QY 439 -----TCCCGAGCTGTTCAACCATCTCGCCTCGCCGGGCAT 471
Db 484 GACCTCTGCGGATCTGTTCTTCCTCGTCTCCAGCTGTTTCAACCATCTCGCCTCGCCGGGCAT 543
QY 472 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 531
Db 544 GAGACGGTGCAGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATG 603
QY 532 GCTTGGGATATGATGAACTGGT 556
Db 604 GCTTGGGATATGATGAACTGGT 628

RESULT 13

US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/612,973
 FILING DATE: 11-MAR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4100
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2433 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2430
 NAME/KEY: mat_peptide
 LOCATION: 1..2427
 US-08-612-973-49

Query Match 84.0%; Score 471; DB 3; Length 2433;
 Best Local Similarity 88.5%; Pred. No. 2.5e-117;
 Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
 4 TTGGTAAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGGGGTACATTCG 63
 355 TTGGTAAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGGGGTACATTCG 414
 64 CTCGTCGCGGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 123
 415 CTCGTCGCGGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 474
 124 GAGGACGGCGTGAATATGCAACAGGGAATTGGCGGCTTCCTTCTCTATCTCTCTC 183
 475 GAGGACGGCGTGAATATGCAACAGGGAATTGGCGGCTTCCTTCTCTATCTCTCTC 534
 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
 535 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
 244 GGGATGTACCATGTCAAGCAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGGAC 303
 595 GGGATGTACCATGTCAAGCAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGGAC 654
 304 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
 655 ATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
 364 TGGGTAGCGCTACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
 715 TGGGTAGCGCTACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
 424 CGAGCCACGTCGAT----- 438
 775 CGAGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834
 439 -----TCCAGCTGTTTCCACATCTCGGCTCGCGGCGCAT 471
 835 GACCTCTCGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
 472 GAGACGGTGCAGGACTGCAATTGCTCAATTATCTCCCGGCCACATAAGCGGTACCGGTATG 531

895 GAGACGGTGCAGGACTGCAATTGCTCAATCTCCCGCCACATACCGGTCACCGTATG 954
 532 GCTTGGGATATGATGATGAACGCT 556
 955 GCTTGGGATATGATGATGAACGCT 979
 RESULT 14
 US-08-927-597-49
 ; Sequence 49, Application US/08927597
 ; Patent No. 6245503
 ; GENERAL INFORMATION:
 ; APPLICANT: MAERTENS, GEERT
 ; APPLICANT: BOSMAN, FONS
 ; APPLICANT: DE MARTYNOFF, GUY
 ; APPLICANT: BUYSE, MARIE-ANGE
 ; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
 ; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
 ; NUMBER OF SEQUENCES: 111
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHVE P.C.
 ; STREET: 1100 NORTH GLEBE ROAD
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/927,597
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/612,973
 ; FILING DATE: 11-MAR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BYRNE, THOMAS E.
 ; REGISTRATION NUMBER: 32,205
 ; REFERENCE/DOCKET NUMBER: 1487-10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 816-4100
 ; TELEFAX: (703) 816-4100
 ; INFORMATION FOR SEQ ID NO: 49:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2433 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..2430
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 1..2427
 ; US-08-927-597-49
 Query Match 84.0%; Score 471; DB 3; Length 2433;
 Best Local Similarity 88.5%; Pred. No. 2.5e-117;
 Matches 553; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
 4 TTGGTAAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGGGGTACATTCG 63
 355 TTGGTAAAGGTCATGATACCTTACATGCGGCTTCGCGACCTCGGGGTACATTCG 414
 64 CTCGTCGCGGCCCCCTAGGGGGCGCTGCCAGGCGCTTCGCGCATGGCGTCCGGTTCG 123

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2294.49 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-25
Perfect score: 606
Sequence: 1 ATGTTGGTAAGTCAATCGA.....TGCTCCGATCTCTTAATAG 606

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	606	100.0	606	6	A48687	A48687 Sequence 25
2	606	100.0	606	6	AR157338	AR157338 Sequence
3	606	100.0	606	6	AX452774	AX452774 Sequence
4	606	100.0	606	6	AX685026	AX685026 Sequence
5	598.2	98.7	723	6	A48683	A48683 Sequence 21
6	598.2	98.7	723	6	AR157336	AR157336 Sequence
7	598.2	98.7	723	6	AX452770	AX452770 Sequence
8	598.2	98.7	723	6	AX685022	AX685022 Sequence
9	597.4	98.6	636	6	A48689	A48689 Sequence 27
10	597.4	98.6	636	6	AR157339	AR157339 Sequence
11	597.4	98.6	636	6	AX452776	AX452776 Sequence
12	597.4	98.6	636	6	AX685028	AX685028 Sequence
13	556	91.7	561	6	A48685	A48685 Sequence 23
14	556	91.7	561	6	AR157337	AR157337 Sequence
15	556	91.7	561	6	AX452772	AX452772 Sequence
16	556	91.7	561	6	AX685024	AX685024 Sequence
17	516.2	85.2	795	6	A48667	A48667 Sequence 5
18	516.2	85.2	795	6	AR157325	AR157325 Sequence
19	516.2	85.2	795	6	AX452754	AX452754 Sequence
20	516.2	85.2	795	6	AX685006	AX685006 Sequence
21	513.2	84.7	2082	6	A48709	A48709 Sequence 47
22	513.2	84.7	2082	6	AR157350	AR157350 Sequence
23	513.2	84.7	2082	6	AX452796	AX452796 Sequence
24	513.2	84.7	2082	6	AX685048	AX685048 Sequence
25	513.2	84.7	2433	6	A48711	A48711 Sequence 49
26	513.2	84.7	2433	6	AR157351	AR157351 Sequence
27	513.2	84.7	2433	6	AX452798	AX452798 Sequence
28	513.2	84.7	2433	6	AX685050	AX685050 Sequence
29	460.4	76.0	3296	14	AB008441	AB008441 Hepatitis
30	460.4	76.0	3296	14	AB008442	AB008442 Hepatitis
31	460.4	76.0	3296	14	AB008443	AB008443 Hepatitis
32	460.4	76.0	9379	14	AF207766	AF207766 Hepatitis
33	457.2	75.4	9379	14	AF165052	AF165052 Hepatitis
34	457.2	75.4	9418	14	HCV132996	HCV132996 Hepatitis
35	456.4	75.3	633	6	A48669	A48669 Sequence 7
36	456.4	75.3	633	6	AR157326	AR157326 Sequence
37	456.4	75.3	633	6	AX452756	AX452756 Sequence
38	456.4	75.3	633	6	AX685008	AX685008 Sequence
39	455.6	75.2	3296	14	AB008445	AB008445 Hepatitis
40	455.6	75.2	3296	14	AB008446	AB008446 Hepatitis
41	455.6	75.2	9344	14	AB049096	AB049096 Hepatitis
42	455.6	75.2	9379	14	HCVPOLYP	AJ000009 Hepatitis
43	455.6	75.2	9386	14	AF165056	AF165056 Hepatitis
44	454	74.9	1615	14	HPCNS1SRJ	M74813 Hepatitis C
45	454	74.9	3296	14	AB008447	AB008447 Hepatitis

ALIGNMENTS

RESULT 1
A48687
LOCUS Sequence 25 from Patent WO9604385.
DEFINITION A48687
ACCESSION A48687
VERSION A48687.1 GI:2302400
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 606)
AUTHORS Maertens,G., Boeman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 25 15-FEB-1996;

Mon Dec 22 13:28:41 2003

LOCUS		AR157338	606 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION		Sequence 25 from patent US 6245503.				
ACCESSION		AR157338				
VERSION		AR157338.1	GI:16218271			
KEYWORDS		Unknown.				
SOURCE		Unknown.				
ORGANISM		Unclassified.				
REFERENCE		1 (bases 1 to 606)				
AUTHORS		Maertens, G., Bosman, F., De Martynoff, G. and Buysse, M.-A.				
TITLE		Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use				
JOURNAL		Patent: US 6245503-A 25 12-JUN-2001;				
FEATURES		Location/Qualifiers				
source		1..606				
BASE COUNT		109 a	193 c	167 g	137 t	
ORIGIN		/organism="unknown"				
Query Match		100.0%; Score 606; DB 6; Length 606;				
Best Local Similarity		100.0%; Pred. No. 1.3e-123;				
Matches		606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCCCGACCTCGTGGGGTACATT	60			
DB	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCCCGACCTCGTGGGGTACATT	60			
QY	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGCTT	120			
DB	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGCTT	120			
QY	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTCTCTTCTATCTTC	180			
DB	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTCTCTTCTATCTTC	180			
QY	181	CTCTTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTG	240			
DB	181	CTCTTGGCTTTGCTCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTCGCAACGTG	240			
QY	241	TCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG	300			
DB	241	TCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG	300			
QY	301	GACATGATCATGCACACCCCGGGTGGTCCCTCGGTCGGGAGAACAACTCTTCCCGC	360			
DB	301	GACATGATCATGCACACCCCGGGTGGTCCCTCGGTCGGGAGAACAACTCTTCCCGC	360			
QY	361	TGCTGGGTAGCGCTCACCCCGGTCGCTCGAGCTAGGAAAGCCAGCGTCCCCACGACA	420			
DB	361	TGCTGGGTAGCGCTCACCCCGGTCGCTCGAGCTAGGAAAGCCAGCGTCCCCACGACA	420			
QY	421	ATACGACGCGCTCAGATTCACAGCTGTTCCAGCTTCCCTCGCCGTCATGAGCGGTG	480			
DB	421	ATACGACGCGCTCAGATTCACAGCTGTTCCAGCTTCCCTCGCCGTCATGAGCGGTG	480			
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGTCACATAACGGGTACCGTATGGCTTGGGAT	540			
DB	481	CAGGACTGCAATTTGCTCAATCTATCCCGGTCACATAACGGGTACCGTATGGCTTGGGAT	540			
QY	541	ATGATGATGAATGCTCAATCTATCCCGGTCACATAACGGGTACCGTATGGCTTGGGAT	600			
DB	541	ATGATGATGAATGCTCAATCTATCCCGGTCACATAACGGGTACCGTATGGCTTGGGAT	600			
QY	601	TAATAG 606				
DB	601	TAATAG 606				
RESULT 3		AX452774	606 bp	DNA	linear	PAT 06-JUL-2001
LOCUS		AX452774				
DEFINITION		Sequence 25 from Patent EP1211315.				
ACCESSION		AX452774				

VERSION AX452774.1 GI:21712459
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1
AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 25 05-JUN-2002;
Innogenetics N.V. (BE)

FEATURES
source
1. .606
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1. .603
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD38074.1"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRLDGG
VNYATGNLPCGCSFIFLLALLSCLTVPASAYEVNVMVHTNDCNSGSIYVEAADM
IMHTPCVPCVRENNSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSYFGHITGHRMNDMMNWSPTTALVVSQLLRIL"
mat_peptide 1. .600
/product="unnamed"
BASE COUNT 109 a 193 c 167 g 137 t
ORIGIN

Query Match 100.0%; Score 606; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.3e-123;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60

QY 61 CCCTCGTTCGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTT 120
DB 61 CCCTCGTTCGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTT 120

QY 121 CTGGAGGACGGGTGAACATATGCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAAGTG 240

QY 241 TCCGGATGTPACCATGTACGAAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300
DB 241 TCCGGATGTPACCATGTACGAAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300

QY 301 GACATGATATGACACACCCCGGTTGCGTCCCTGCGTTTCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATATGACACACCCCGGTTGCGTCCCTGCGTTTCGGAGAACAACTCTTCCCGC 360

QY 361 TGGTGGTGGTTCACCCCGGCTCGAGCTAGGAAACGCGTCCCGACACGACA 420
DB 361 TGGTGGTGGTTCACCCCGGCTCGAGCTAGGAAACGCGTCCCGACACGACA 420

QY 601 TAATAG 606
DB 601 TAATAG 606

RESULT 4
AX685026
LOCUS
DEFINITION Sequence 25 from Patent WO0205548.
ACCESSION AX685026
VERSION AX685026.1 GI:29371431
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1
AUTHORS Maertens, G., Bosman, F. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: WO 0205548-A 25 18-JUL-2002;
INNOCENETICS N.V. (BE)

FEATURES
source
1. .606
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1. .603
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD86527.1"
/db_xref="GI:29371432"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGAAALAHGVRLDGG
VNYATGNLPCGCSFIFLLALLSCLTVPASAYEVNVMVHTNDCNSGSIYVEAADM
IMHTPCVPCVRENNSSRCWVALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSYFGHITGHRMNDMMNWSPTTALVVSQLLRIL"
mat_peptide 1. .600
/product="unnamed"
BASE COUNT 109 a 193 c 167 g 137 t
ORIGIN

Query Match 100.0%; Score 606; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.3e-123;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCACCTCGTGGGTACATT 60

QY 61 CCCTCGTTCGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTT 120
DB 61 CCCTCGTTCGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGTT 120

QY 121 CTGGAGGACGGGTGAACATATGCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACATATGCAACAGGAAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCGAAGTG 240

QY 241 TCCGGATGTPACCATGTACGAAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300
DB 241 TCCGGATGTPACCATGTACGAAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGC 300

QY 301 GACATGATATGACACACCCCGGTTGCGTCCCTGCGTTTCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATATGACACACCCCGGTTGCGTCCCTGCGTTTCGGAGAACAACTCTTCCCGC 360

QY 361 TGGTGGTGGTTCACCCCGGCTCGAGCTAGGAAACGCGTCCCGACACGACA 420
DB 361 TGGTGGTGGTTCACCCCGGCTCGAGCTAGGAAACGCGTCCCGACACGACA 420

421	ATACGACGCCACGTCGATTCCACAGCTGTTACCACTCTCGCTCGCCGCGCATGAGACGGTG	480
421	ATACGACGCCACGTCGATTCCACAGCTGTTACCACTCTCGCTCGCCGCGCATGAGACGGTG	480
481	CAGGACTCGCAATTGGTCCAATCTATCCCGGCCACACATAACCGGTCACCGTATGGCTTGGGAT	540
481	CAGGACTCGCAATTGGTCCAATCTATCCCGGCCACACATAACCGGTCACCGTATGGCTTGGGAT	540
541	ATGATGATGAACCTGCTCGCTTACACGGCCCTTGTTGTTATCGCAGCTGCTCCGGATCCTC	600
541	ATGATGATGAACCTGCTCGCTTACACGGCCCTTGTTGTTATCGCAGCTGCTCCGGATCCTC	600
601	TAATAG 606	
601	TAATAG 606	

RESULT 5			
LOCUS	A48683	723 bp	DNA
DEFINITION	Sequence 21 from Patent WO9604385.		
ACCESSION	A48683		
VERSION	A48683.1	GI:2302396	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 723)		

FILE	THERAPEUTIC USE									
JOURNAL	Patent: WO 9604385-A 21 15-FEB-1996; INNOGENETICS NV (BE) Other publication CA 2172273 960215 Other publication AU 3382495 960304. Location/Qualifiers									
FEATURES	1..723 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644" 1..720 /note="unnamed protein product" /codon_start=1 /protein_id="CAA03134.1" /db_xref="GI:2302397" /translation="MLGKVIDTLTCGFADLVGVIPLVAGPLGAPARALAHGVRLVEDG VNYATNLPGCSFISFILALLSCLTPASAVEYNNVGMVHTNDCSNLSIVYEADIM IMHTPGCVPCVRENNNSRCWVALPTLAAENASVPTTIPRRHVDLSOLFTISPRRHETV QDCNCSIYPGHIITGHRWADMMNNVSTTALNVSVQLLRIPQAVVDMVAGAHGVLVLAG AAYSMVGNWAKVLIVMLIFAP" mat_peptide 1..717 /product="unnamed"									
BASE COUNT	126 a	220 c	208 g	169 t						
ORIGIN										
Query Match	98.7%; Score 598.2; DB 6; Length 723;									
Best Local Similarity	99.5%; Pred. No. 6.8e-122;									
Matches	600; Conservative	0; Mismatches 3;				Indels	0;	Gaps	0;	
Qy	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATT	60							
Db	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATT	60							
Qy	61	CCGCTCGTCGCGCCCCCTAGGGGCGCGTCGACGGCCCTGGCGCATGGGTCGGGGTT	120							
Db	61	CCGCTCGTCGCGCCCCCTAGGGGCGCGTCGACGGCCCTGGCGCATGGGTCGGGGTT	120							
Qy	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC	180							
Db	121	CTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCGGTGTCTTTCTCTATCTTC	180							

181 CACCTGGCCTTTGCTGTCTGTCTGACCGTTCAGCTTCGGCTATGAACTGGCAACGTG 240

```

QY 241 TCCGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGG 300
DB 241 TCCGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTACACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCAACGACA 420
DB 361 TGTGGGTAGCGCTACACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCAACGACA 420
QY 421 ATACGACGCCACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
DB 421 ATACGACGCCACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACCGGTCAACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACCGGTCAACCGTATGGCTTGGAT 540
QY 541 ATGATGATGAACCTGCTCGCTCAACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
DB 541 ATGATGATGAACCTGCTCGCTCAACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
QY 601 TAA 603
DB 601 CAA 603

RESULT 7
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
          Viruses; serNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, P., de Martynoff, G. and Buyse, M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
  source
    1..723
    /organism="Hepatitis C virus"
    /mol_type="genomic DNA"
    /db_xref="taxon:11103"
  CDS
    1..720
    /notes="unnamed protein product"
    /codon_start=1
    /protein_id="CAD38072.1"
    /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLBDG
    VNYATGNLPCGCSFIFLLALLSCLTVPASAYEVNRYGMVHTNDSCNSIYVEAADM
    IMHTPCVPCVRNNSRCWVALTPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
    ODCNCSIYPGHITGHMAMDMNWSPTTALVVSQLLRIPQAVVDVMAHGVGLAGL
    AYSVMGVNAKVLIVMLLPAP"
  mat_peptide
    1..717
    /product="unnamed"
    /base_count
      126 a 220 c 208 g 169 t
  ORIGIN
    Query Match 98.7%; Score 598.2; DB 6; Length 723;
    Best Local Similarity 99.5%; Pred. No. 6.8e-122;
    Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCACTGATCCCTACATCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCACTGATCCCTACATCGGCTTCGCCGACCTCGTGGGTACATT 60

```

```

QY 61 CCCTCGTGGCGCCGCCCTAGGGGCGCTGCCAGGCGCCTGCGCATGCGTCCGGTT 120
DB 61 CCCTCGTGGCGCCGCCCTAGGGGCGCTGCCAGGCGCCTGCGCATGCGTCCGGTT 120
QY 121 CTGAGGAGCGGCGTGAACCTATGCAACAGAGAAATTTCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGAGGAGCGGCGTGAACCTATGCAACAGAGAAATTTCCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGTCTGACCGTTCCAGCTTTCCGCTTATGAAGTGCACAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGTCTGACCGTTCCAGCTTTCCGCTTATGAAGTGCACAAGTG 240
QY 241 TCCGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGG 300
DB 241 TCCGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGACGG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGCTCGGAGAACAACTCTTCCCGC 360
QY 361 TGTGGGTAGCGCTCAACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCAACGACA 420
DB 361 TGTGGGTAGCGCTCAACCCCGACGCTCGCAGCTAGGAAACCCAGCGTCCCAACGACA 420
QY 421 ATACGACGCCACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
DB 421 ATACGACGCCACGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGCATGAGCGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACCGGTCAACCGTATGGCTTGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCACATAACCGGTCAACCGTATGGCTTGGAT 540
QY 541 ATGATGATGAACCTGCTCGCTCAACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
DB 541 ATGATGATGAACCTGCTCGCTCAACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCTC 600
QY 601 TAA 603
DB 601 CAA 603

RESULT 8
AX685022
LOCUS AX685022 723 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 21 from Patent WO02055548.
ACCESSION AX685022
VERSION AX685022.1 GI:29371427
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
          Viruses; serNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus.
REFERENCE 1
AUTHORS Maertens, G., Bosman, P. and Buyse, M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
          therapeutic use
JOURNAL Patent: WO 02055548-A 21 18-JUL-2002;
Innogenetics N.V. (BE)
FEATURES
  source
    1..723
    /organism="Hepatitis C virus"
    /mol_type="genomic DNA"
    /db_xref="taxon:11103"
  CDS
    1..720
    /notes="unnamed protein product"
    /codon_start=1
    /protein_id="CAD86525.1"
    /translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLBDG
    VNYATGNLPCGCSFIFLLALLSCLTVPASAYEVNRYGMVHTNDSCNSIYVEAADM
    IMHTPCVPCVRNNSRCWVALTPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
    ODCNCSIYPGHITGHMAMDMNWSPTTALVVSQLLRIPQAVVDVMAHGVGLAGL
    AYSVMGVNAKVLIVMLLPAP"

```

source	1. .636										
	/organism="unidentified"										
	/mol_type="genomic DNA"										
	/db_xref="taxon:32644"										
CDS	1. .633										
	/note="unnamed protein product"										
	/codon_start=1										
	/protein_id="CAA03137.1"										
	/db_xref="GI:2302403"										
	/translation="MLKVIYDTLCGFADLVGYIPLVGA PLGGAAALAHGVRVLIED										
	VNYATGNLPGGSFSLFLIALSLCTVPASAYEVNUSGMVHTNDCSSSIVVEAADM										
	IMHTPGCVPCVRENSSRCWALPTTLAARNASVPTTIRRHVDLSOLFTLSPRRHETV										
	ODCNSIYPGHITGHRAWDMNMWSPPTALVWSQLLRI VIEGRHHHHH"										
mat_peptide	1. .630										
	/product="unnamed"										
BASE COUNT	119 a	203 c	174 g	140 t							
ORIGIN											
	Query Match	98.6%;	Score 597.4;	DB 6;	Length 636;						
	Best Local Similarity	99.8%;	Pred. No. 1e-121;								
	Matches 598;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;					
Qy	1	ATGTTGGTAAAGTCAATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60								
Db	1	ATGTTGGTAAAGTCAATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60								
Qy	61	CCGCTCGTGGGGCCCCCTAGGGGGCGGTGCCAGGGCCCTGGCGCATGGCGTCCGGTT	120								
Db	61	CCGCTCGTGGGGCCCCCTAGGGGGCGGTGCCAGGGCCCTGGCGCATGGCGTCCGGTT	120								
Qy	121	CTGGAGGACGGGTGAACATATGCAACAGGAAATTTGCCCGGTGCTCTTTCTCTATCTTC	180								
Db	121	CTGGAGGACGGGTGAACATATGCAACAGGAAATTTGCCCGGTGCTCTTTCTCTATCTTC	180								
Qy	181	CTCTTGGCTTTGTGTCCTGTCTGACCGTTCCAGCTTCGCGTTATGAAGTGCACACGTG	240								
Db	181	CTCTTGGCTTTGTGTCCTGTCTGACCGTTCCAGCTTCGCGTTATGAAGTGCACACGTG	240								
Qy	241	TCCGGGATGTACATGTCAGAACGACTGCTCAACTCAAGCATTTGTATGAGGCAGCG	300								
Db	241	TCCGGGATGTACCATGTGACAAACGACTGCTCAACTCAAGCATTTGTATGAGGCAGCG	300								
Qy	301	GACATGATCATGACACCCCCGGGTGCGTCCCTGGCTTCGGAGAACAACTTTCCCGG	360								
Db	301	GACATGATCATGACACCCCCGGGTGCGTCCCTGGCTTCGGAGAACAACTTTCCCGG	360								
Qy	361	TGCTGGGTAGCGGTCAACCCCAACGCTCGAGCTAGGAACGCCAGCGTCCCCACACGACA	420								
Db	361	TGCTGGGTAGCGGTCAACCCCAACGCTCGAGCTAGGAACGCCAGCGTCCCCACACGACA	420								
Qy	421	ATACGAGCCACGTCGATTCACCATCTGTTCAACATCTCGCCTCGCGGCATGAGACGGTG	480								
Db	421	ATACGAGCCACGTCGATTCACCATCTGTTCAACATCTCGCCTCGCGGCATGAGACGGTG	480								
Qy	481	CAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTACACGATGGCTTGGGAT	540								
Db	481	CAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTACACGATGGCTTGGGAT	540								
Qy	541	ATGATGATGAACGTGTCGCTTACAAACGSCCTGCTGGTATCGCAGCTGCTCCGGATCCT	599								
Db	541	ATGATGATGAACGTGTCGCTTACAAACGSCCTGCTGGTATCGCAGCTGCTCCGGATCCT	599								

RESULT 10				
AR157339				
LOCUS	AR157339		636 bp	DNA
DEFINITION	Sequence 27 from patent US 6245503.			
ACCESSION	AR157339			
VERSION	AR157339.1		GI:16218273	
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
	Unclassified.			
				linear
				PAT 17-OCT-2001

mat_peptide		1. .717		/product="unnamed"			
BASE COUNT	126 a	220 c	208 g	169 t			
ORIGIN							
Query Match	98.7%	Score 598.2;	DB 6;	Length 723;			
Best Local Similarity	99.5%	Pred. No. 6.8e-122;					
Matches 600;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;			
Qy	1	ATGTTGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCAGACCTCGTGGGGTACATT	60				
Db	1	ATGTTGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGCCAGACCTCGTGGGGTACATT	60				
Qy	61	CGGCTGTCGGCGCCCCCTTAGGGGCGCTCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120				
Db	61	CGGCTGTCGGCGCCCCCTTAGGGGCGCTCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120				
Qy	121	CTGAGAGACGCGTGAATCTATGCAACAGGAAATTTGCCGGTTGCTCTTTCTCTATCTTC	180				
Db	121	CTGAGAGACGCGTGAATCTATGCAACAGGAAATTTGCCGGTTGCTCTTTCTCTATCTTC	180				
Qy	181	CTCTTGGCTTTGCTGTCTGTCTCAACCGTTCCAGTTCGCGTTATGAAGTGCACACGTG	240				
Db	181	CTCTTGGCTTTGCTGTCTGTCTCAACCGTTCCAGTTCGCGTTATGAAGTGCACACGTG	240				
Qy	241	TCCGGGATGTACCATGTACAAACGACTGCTCCAACTCAAGCATTTGTATGAGGCAGCG	300				
Db	241	TCCGGGATGTACCATGTACAAACGACTGCTCCAACTCAAGCATTTGTATGAGGCAGCG	300				
Qy	301	GACATGATCATGCACACCCCCGGGTGCGTTCGCTTCGGGAGAACTCTTTCGGC	360				
Db	301	GACATGATCATGCACACCCCCGGGTGCGTTCGCTTCGGGAGAACTCTTTCGGC	360				
Qy	361	TGCTGGGTAGCGCTCACCCCAACGTCGAGTAGGAACGCCAGGTCCCAACACGACA	420				
Db	361	TGCTGGGTAGCGCTCACCCCAACGTCGAGTAGGAACGCCAGGTCCCAACACGACA	420				
Qy	421	ATACAGCGCACGTGGATTCAGCTGTTTCAACATCTCGCTTCGGCGCATGAGACGGTG	480				
Db	421	ATACAGCGCACGTGGATTCAGCTGTTTCAACATCTCGCTTCGGCGCATGAGACGGTG	480				
Qy	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTTCGGTTGGAT	540				
Db	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACGGGTTCGGTTGGAT	540				
Qy	541	ATGATGATGAATGGTTCGCTTCAACGGGCCCTGGTGGTATCGACGTGTCGGATCCCT	600				
Db	541	ATGATGATGAATGGTTCGCTTCAACGGGCCCTGGTGGTATCGACGTGTCGGATCCCA	600				
Qy	601	TAA 603					
Db	601	CAA 603					

RESULT	9
LOCUS	A48689
DEFINITION	Sequence 27 from Patent WO9604385.
ACCESSION	A48689
VERSION	A48689.1 GI:2302402
KEYWORDS	.
SOURCE	unidentified
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 636)
TITLE	Maertens,G., Bosman,F., De,M.G. and Buysse,M. PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL	Patent: WO 9604385-A 27 15-FEB-1996; INNOCENTICS NV (BE)
COMMENT	Other publication CA 2172273 960215 Other publication AU 3382495 960304.
FEATURES	Location/Qualifiers

REFERENCE 1 (bases 1 to 636)
AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 27 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..636
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN /organism="unknown"
Query Match 98.6%; Score 597.4; DB 6; Length 636;
Best Local Similarity 99.8%; Pred. No. 1e-121; Indels 0; Gaps 0;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGTCCGGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGTCCGGGTT 120
QY 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTGACCGCTTCAGCTTCGGCTTATGAAGTGCACAAACGTTG 240
Db 181 CTCTTGGCTTTGCTGTCCTGTGACCGCTTCAGCTTCGGCTTATGAAGTGCACAAACGTTG 240
QY 241 TCGGGATGATACATGTCACGAACGACTGCTCCAACTCAAGCAATTTGTTATGAGGACGCG 300
Db 241 TCGGGATGATACATGTCACGAACGACTGCTCCAACTCAAGCAATTTGTTATGAGGACGCG 300
QY 421 ATACGACGCGCATCGATTCCAGCTGTTCAACATTCGCTCCGCGCATGAGACGGTG 480
Db 421 ATACGACGCGCATCGATTCCAGCTGTTCAACATTCGCTCCGCGCATGAGACGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGCTCGCTACACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCT 599
Db 541 ATGATGATGAACCTGCTCGCTACACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCT 599
RESULT 11
AX452776
LOCUS AX452776 636 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 27 from Patent EP1211315.
ACCESSION AX452776
VERSION AX452776.1 GI:21712461
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 27 05-JUN-2002;
FEATURES Innogenetics N.V. (BE)
Location/Qualifiers

source 1..636
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
CDS 1..633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38075.1"
/db_xref="GI:21712462"
/translation="MLGKVIDLTTCGFADLVGYPLVGAPLGGAAALAHGVRLBDE
VNVATGNLPGCSFSLFLLALLSCLTVPASAVEVRNVGMVHTNDCNSISIVYEADM
INHPTGCVPCVRENSSRCWVALFTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNSIYFEHITGHRMWMNMNWSPTTALVSQLLRIVIEGRHHHHH"
mat_peptide 1..630
/product="unnamed"
BASE COUNT 119 a 203 c 174 g 140 t
ORIGIN
Query Match 98.6%; Score 597.4; DB 6; Length 636;
Best Local Similarity 99.8%; Pred. No. 1e-121;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGTCCGGGTT 120
Db 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGCGCATGCGCGTCCGGGTT 120
QY 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGTCCTGTGACCGCTTCAGCTTCGGCTTATGAAGTGCACAAACGTTG 240
Db 181 CTCTTGGCTTTGCTGTCCTGTGACCGCTTCAGCTTCGGCTTATGAAGTGCACAAACGTTG 240
QY 241 TCGGGATGATACATGTCACGAACGACTGCTCCAACTCAAGCAATTTGTTATGAGGACGCG 300
Db 241 TCGGGATGATACATGTCACGAACGACTGCTCCAACTCAAGCAATTTGTTATGAGGACGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTTCGCTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTTCGCTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGTAGCGCTACACCCCGGCTCGCAGCTAGGAAACCCAGCGTCCCCACACGACA 420
Db 361 TGCTGGTAGCGCTACACCCCGGCTCGCAGCTAGGAAACCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCGCATCGATTCCAGCTGTTCAACATTCGCTCCGCGCATGAGACGGTG 480
Db 421 ATACGACGCGCATCGATTCCAGCTGTTCAACATTCGCTCCGCGCATGAGACGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGCTCGCTACACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCT 599
Db 541 ATGATGATGAACCTGCTCGCTACACGCGCCCTGGTGTATCGCAGCTGCTCCGGATCCT 599
RESULT 12
AX685028
LOCUS AX685028 636 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 27 from Patent WO0205548.
ACCESSION AX685028
VERSION AX685028.1 GI:29371433
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;

Mon Dec 22 13:28:41 2003

LOCUS	A48685	Sequence 23 from Patent WO9604385.	561 bp	DNA	linear	PAT 07-MAR-1997
DEFINITION	A48685					
ACCESSION	A48685					
VERSION	A48685.1	GI:2302398				
KEYWORDS						
SOURCE	unidentified					
ORGANISM	unidentified					
REFERENCE	1	(bases 1 to 561)				
AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buyse,M.					
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE					
JOURNAL	PATENT: WO 9604385-A 23 15-FEB-1996;					
COMMENT	INNOGENETICS NV (BE)					
Other publication	CA 2172273 960215					
Other publication	AU 3382495 960304.					
FEATURES	Location/Qualifiers					
source	1..561	/organism="unidentified"				
		/mol_type="genomic DNA"				
	1..558	/db_xref="taxon:32644"				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAA03135.1"				
		/db_xref="GI:2302398"				
		/translation="MLGKVIDTLTCGFADLVGYIPLVGAIPLGGAARALAHGVRLDGG VNYATGNLPGCSFSIFLLALLSLCTVPASAYEVNVMGMYHVNDCSNSSIVYEADM IMHTPGCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDSQLFTISPRRHETV QDCNCSYFGHITGHRMADMMNMW"				
CDS	1..555	/product="unnamed"				
mat_peptide	103 a	176 c	155 g	127 t		
BASE COUNT						
ORIGIN						
Query Match	91.7%;	Score 556;	DB 6;	Length 561;		
Best Local Similarity	100.0%;	Pred. No. 1.4e-112;				
Matches	556;	Conservative	0;	Mismatches	0;	Gaps
QY	1	ATGTTGGGTAAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60			
Db	1	ATGTTGGGTAAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60			
QY	61	CGCTCGTGGGGCCCCCTAGGGGCGCTGCCAGGCGCTTCCGCGACCTCGTGGGGTACATT	120			
Db	61	CGCTCGTGGGGCCCCCTAGGGGCGCTGCCAGGCGCTTCCGCGACCTCGTGGGGTACATT	120			
QY	121	CTGGAGGACGGGCGTGAACCTATGCAACAGGGAAATTTGCCGGTTCCTTTCTCTATCTTC	180			
Db	121	CTGGAGGACGGGCGTGAACCTATGCAACAGGGAAATTTGCCGGTTCCTTTCTCTATCTTC	180			
QY	181	CTCTTGGCTTTCGCTGCTGATGCAACAGGGAAATTTGCCGGTTCCTTTCTCTATCTTC	240			
Db	181	CTCTTGGCTTTCGCTGCTGATGCAACAGGGAAATTTGCCGGTTCCTTTCTCTATCTTC	240			
QY	241	TCCGGGATGACCATGTGCAAGCGACTGCTCCAACTCAAGCATTTGTATGAGGACGC	300			
Db	241	TCCGGGATGACCATGTGCAAGCGACTGCTCCAACTCAAGCATTTGTATGAGGACGC	300			
QY	301	GACATGATCATGCACACCCCGGGTGGTCCCTTGGTTCCGGAGAACAACTCTTCCCGC	360			
Db	301	GACATGATCATGCACACCCCGGGTGGTCCCTTGGTTCCGGAGAACAACTCTTCCCGC	360			
QY	361	TGCTGGGTAGGCTCAACCTATCCCGGCAATACAGGCTTCCCGGATGGGAT	420			
Db	361	TGCTGGGTAGGCTCAACCTATCCCGGCAATACAGGCTTCCCGGATGGGAT	420			
QY	421	ATAGATGTAACCTGGTCCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCT	480			
Db	421	ATAGATGTAACCTGGTCCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCT	480			
QY	481	CAGGACTGCAATGCTCAATCTATCCCGGCAATACAGGCTTCCCGGATGGGAT	540			
Db	481	CAGGACTGCAATGCTCAATCTATCCCGGCAATACAGGCTTCCCGGATGGGAT	540			
QY	541	ATGATGATGAACTGGTCCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCT	599			
Db	541	ATGATGATGAACTGGTCCCTTACAAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCT	599			
RESULT 13						
A48685						

Db 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATACGGGTACCGGTATGGCTTGGGAT 540

QY 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 14

LOCUS AR157337

DEFINITION Sequence 23 from patent US 6245503.

ACCESSION AR157337

VERSION AR157337.1 GI:16218270

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 561)

AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.

TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use

JOURNAL Patent: US 6245503-A 23 12-JUN-2001;

FEATURES Location/Qualifiers

source 1..561

BASE COUNT 103 a 176 c 155 g 127 t

ORIGIN

Query Match 91.7%; Score 556; DB 6; Length 561;

Best Local Similarity 100.0%; Pred. No. 1.4e-112;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

Db 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

QY 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGGCGATGCGTCCGGTT 120

Db 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGGCGATGCGTCCGGTT 120

QY 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTCTGACCGTTTCAGCGTTTCGCGTTTATGAAGTGCACAACGTTG 240

Db 181 CTCTTGGCTTTGCTGCTCTGACCGTTTCAGCGTTTCGCGTTTATGAAGTGCACAACGTTG 240

QY 241 TCCGGGATGATCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

Db 241 TCCGGGATGATCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCGGCTCGAGCTAGGAACCGCAGCGTCCCGACGACA 420

Db 361 TGCTGGGTAGCGCTCACCCCGGCTCGAGCTAGGAACCGCAGCGTCCCGACGACA 420

QY 421 ATACGAGGACATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 480

Db 421 ATACGAGGACATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540

Db 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540

QY 541 ATGATGATGAACCTGGT 556

Db 541 ATGATGATGAACCTGGT 556

RESULT 15

AX452772

LOCUS

DEFINITION Sequence 23 from Patent EP1211315.

ACCESSION AX452772

VERSION AX452772.1 GI:21712457

KEYWORDS

SOURCE Hepatitis C virus

ORGANISM Hepatitis C virus

REFERENCE 1

AUTHORS Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.A.

TITLE Recombinant vectors for producing hcv envelope proteins

JOURNAL Patent: EP 1211315-A 23 05-JUN-2002;

INNOCENETICS N.V. (BE)

FEATURES Location/Qualifiers

source 1..561

/organism="Hepatitis C virus"

/mol_type="genomic DNA"

/db_xref="taxon:11103"

1..558

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD38073.1"

/db_xref="GI:21712458"

/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGLGAARALAHGVRVLEBDG VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNRVNGMYHTVNDGNSNIIIVYEADM IMHTPGCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDSQLFTISPRRHETV ODCNSIYPGHITGHRMAWDMNMW"

mat_peptide 1..555

/product="unnamed"

BASE COUNT 103 a 176 c 155 g 127 t

ORIGIN

Query Match 91.7%; Score 556; DB 6; Length 561;

Best Local Similarity 100.0%; Pred. No. 1.4e-112;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

Db 1 ATGTTGGGTAAGTTCATCGATACCCCTTACATGCGGCTTGCCGACCTCGTGGGGTACATT 60

QY 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGGCGATGCGTCCGGTT 120

Db 61 CCGCTCGTGGGCGCCCGCTAGGGGGCGTCCAGGGCCCTGGCGATGCGTCCGGTT 120

QY 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

Db 121 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTCTGACCGTTTCAGCGTTTCGCGTTTATGAAGTGCACAACGTTG 240

Db 181 CTCTTGGCTTTGCTGCTCTGACCGTTTCAGCGTTTCGCGTTTATGAAGTGCACAACGTTG 240

QY 241 TCCGGGATGATCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

Db 241 TCCGGGATGATCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 360

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCGGCTCGAGCTAGGAACCGCAGCGTCCCGACGACA 420

Db 361 TGCTGGGTAGCGCTCACCCCGGCTCGAGCTAGGAACCGCAGCGTCCCGACGACA 420

QY 421 ATACGAGGACATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 480

Db 421 ATACGAGGACATGACACACCCCGGGTGGTCCCTTGGGAGAACAACTCTTCCCGC 480

QY 481 CAGGACTGCAATTGCTCAATCTATCCGGCCACATACCGGTCACCGTATGGCTTGGGAT 540

us-09-899-303a-25.rge

Mon Dec 22 13:28:41 2003

Db 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGAT 540
Qy 541 ATGATGATGAAGTGGT 556
Db 541 ATGATGATGAAGTGGT 556

Search completed: December 20, 2003, 02:02:03
Job time : 2295.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 167.534 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-25
Perfect score: 606
Sequence: 1 AGRITGGTAAAGTCAATCA.....TGCTCGGATCCTCTAATAG 606

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	100.0	606	17	AAT12963 HCV E1 construct H
2	606	100.0	606	24	AAL48927 Hepatitis C virus
3	598.2	98.7	723	17	AAT12961 HCV E1 construct H
4	598.2	98.7	723	24	AAL48925 Hepatitis C virus
5	597.4	98.6	636	17	AAT12964 HCV E1 construct H
6	597.4	98.6	636	24	AAL48928 Hepatitis C virus
7	556	91.7	561	17	AAT12962 HCV E1 construct H
8	556	91.7	561	24	AAL48926 Hepatitis C virus

9	516.2	85.2	795	17	AAT12705 HCV E1 construct H
10	516.2	85.2	795	24	AAL48914 Hepatitis C virus
11	513.2	84.7	2082	24	AAL48939 Hepatitis C virus
12	513.2	84.7	2086	17	AAT12973 HCV E1 construct H
13	513.2	84.7	2433	17	AAT12974 HCV E1 construct H
14	502.2	82.9	2434	24	AAL48940 Hepatitis C virus
15	456.4	75.3	633	17	AAT12706 HCV E1 construct H
16	456.4	75.3	633	24	AAL48915 Hepatitis C virus
17	453	74.8	673	19	ABA03491 Cuticle protein 1
18	452.4	74.7	2187	19	ABA03491 NANBHV hepatitis vir
19	452.4	74.7	2540	15	AQ033753 NANBHV genomic fra
20	452.4	74.7	2540	15	AQ033753 Hepatitis C virus
21	452.4	74.7	9605	24	ABK91411 Hepatitis C virus
22	452.4	74.7	9605	24	ABK91424 Hepatitis C virus
23	452.4	74.7	9605	24	ABK91425 Hepatitis C virus
24	452.4	74.7	9605	24	ABK91426 Hepatitis C virus
25	452.4	74.7	9605	24	ABK91428 Hepatitis C virus
26	452.4	74.7	9605	24	ABK91429 Hepatitis C virus
27	452.4	74.7	9605	24	ABK91430 Hepatitis C virus
28	452.4	74.7	9605	24	ABK91431 Hepatitis C virus
29	452.4	74.7	9605	24	ABK91432 Hepatitis C virus
30	452.4	74.7	9605	24	ABK91433 Hepatitis C virus
31	452.4	74.7	9605	24	ABK91434 Hepatitis C virus
32	452.4	74.7	9608	24	ABK91427 Hepatitis C virus
33	452.4	74.7	11062	24	ABK91427 Hepatitis C virus
34	452.4	74.7	11076	21	AA98965 Hepatitis C virus
35	451.6	74.5	636	17	AAT12709 HCV E1 construct H
36	451.6	74.5	636	24	AAL48918 Hepatitis C virus
37	450.8	74.4	1562	19	AAV60672 Fragment #5 isolat
38	450.8	74.4	1880	13	AAQ24467 NANBHV hepatitis vir
39	450.8	74.4	1953	25	AAU55222 Plasmid pIDKE2 DNA
40	450.8	74.4	2829	19	AAV60673 Fragment #6 isolat
41	449.2	74.1	1251	13	AAQ26981 HCV gene 1. Hepat
42	449.2	74.1	2540	13	AAQ26981 Hepatitis C virus
43	449.2	74.1	3360	17	AAT03677 Hepatitis C genome
44	449.2	74.1	3461	15	AQ640668 Non-A, non-B hepat
45	449.2	74.1	3461	16	AAQ30386 5'UTR/CORE/ENV/NS1

ALIGNMENTS

RESULT 1
AAT12963
ID AAT12963 standard; DNA; 606 BP.

AC AAT12963;

DT 24-SEP-1996 (first entry)

XX HCV E1 construct HCC139.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

XX Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

Query Match	100.0%;	Score 606;	DB 17;	Length 606;
Best Local Similarity	100.0%;	Pred. No. 1.2e-152;		
Matches 606;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGTTGGTAAAGTCATCGATCCCTTACATCGGGCTTCGCCACCTCGTGGGGTACATT	60
DB	1	ATGTTGGTAAAGTCATCGATACCTTTTCAATGGGCTTCGCCACCTCGTGGGGTACATT	60
QY	61	CCGCTCGTCCGGCGCCCCCTTAGGGGGCGCTGCACAGGCCCTTGGGCATGGCGTCCGGGTT	120
DB	61	CCGCTCGTCCGGCGCCCCCTTAGGGGGCGCTGCCAGGCCCTTGGGCATGGCGTCCGGGTT	120
QY	121	CTGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTTCTCTATCTTC	180
DB	121	CTGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTTCTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGTTCCGGTTATGAAGTCGCAACGCTG	240
DB	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGTTCCGGTTATGAAGTCGCAACGCTG	240
QY	241	TCCGGGATGACCATGTACACAAACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGC	300
DB	241	TCCGGGATGATCCATGTACAGNACGACTGCTCCRACTCAAGCATTTGTATGAGCGACGC	300
QY	301	GACATGATCATGCAACCCCGGGTGGTGGCTCCCTGGTTCCGGAGAACACTCTTCCCGC	360
DB	301	GACATGATCATGCAACCCCGGGTGGTGGCTCCCTGGTTCCGGAGAACACTCTTCCCGC	360
QY	361	TGCTGGTAGCGCTCAACCCCAACGCTCGCAGCTAGGAAACGCAGCGTCCCCACACGACA	420
DB	361	TGCTGGTAGCGCTCACCCCAACGCTCGCAGCTAGGAAACGCAGCGTCCCCACACGACA	420
QY	421	ATACGACGCCAGTTCGATTTCCAGCTGTTCCACCTCGCCCTCGCGGCATGAGACGGTG	480
DB	421	ATACGACGCCAGTTCGATTTCCAGCTGTTCCACCTCGCCCTCGCGGCATGAGACGGTG	480
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTGGGAT	540
DB	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGCTGGGAT	540
QY	541	ATGATGATGAACCTGGTCGCTCAACAGGCCCTGGTGTATCGCAGCTGCTCCGGATCTCTC	600
DB	541	ATGATGATGAACCTGGTCGCTCAACAGGCCCTGGTGTATCGCAGCTGCTCCGGATCTCTC	600
QY	601	TAATAG 606	

	Query Match	100.0%	Score 606;	DB 24;	Length 606;
	Best Local Similarity	100.0%;	Pos. No. 1.2e-152;		
	Matches 606;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGTTGGGTAAAGGTCATCGATACCTTACATCGCGCTTCGCGCAACCTCGTGGGGTACATT	60		
Db	1	ATGTTGGGTAAAGGTCATCGATACCTTACATCGCGCTTCGCGCAACCTCGTGGGGTACATT	60		
Qy	61	CGGCTCGTCGGCGCCCCCTCAGGGGGCGCTCGCAGGGCCCTGGCGCATGGGTCGCGGTT	120		
Db	61	CGGCTCGTCGGCGCCCCCTCAGGGGGCGCTCGCAGGGCCCTGGCGCATGGGTCGCGGTT	120		
Qy	121	CTGGAGCAGCGGTGAACCTATGCAACACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180		
Db	121	CTGGAGCAGCGGTGNACTAATGCAACACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180		
Qy	181	CTCTTGGGCTTGTGTCCTGTCTGACCGTTCCAGCTTCGCGTTATGAAGTCGCAACGTG	240		

	Query Match	100.0%	Score 606;	DB 24;	Length 606;
	Best Local Similarity	100.0%;	Pos. No. 1.2e-152;	Indels 0;	Gaps 0;
	Matches 606;	Conservative	0;	Mismatches 0;	
Qy	1	ATGTTGGGTAAAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60		
Db	1	ATGTTGGGTAAAGGTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60		
Qy	61	CGGCTCGTCGGCGCCCCCTCAGGGGGCGCTCGCAGGGCCCTGGCGGCATGGGCTCCGGGTT	120		
Db	61	CGGCTCGTCGGCGCCCCCTCAGGGGGCGCTCGCAGGGCCCTGGCGGCATGGGCTCCGGGTT	120		
Qy	121	CTGGAGCAGCGGTGAACCTATGCAACACAGGGAATTTGCCCGGTTGCTCTTTTCTATATCTTC	180		
Db	121	CTGGAGCAGCGGTGNACTAIGCAACACAGGGAATTTGCCCGGTTGCTCTTTTCTATATCTTC	180		
Qy	181	CTCTTGGGCTTGTGTCCTGTCTGACCGTTCCAGCTTCCGGCTTATGAAGTCGCAACGTG	240		

CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.

XX SQ Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;

Query Match 98.6%; Score 597.4; DB 17; Length 636;
 Best Local Similarity 99.8%; Pred. No. 2.4e-150;
 Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTGATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGGTACATT 60
 DB |||||
 DB 1 ATGTTGGGTAAGGTGATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGGTACATT 60
 QY 61 CGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGCGCATGCGCGTTCGGGTT 120
 DB |||||
 DB 61 CGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGCGCATGCGCGTTCGGGTT 120
 QY 121 CTGGAGGACGGCGTGAATGATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTC 180
 DB |||||
 DB 121 CTGGAGGACGGCGTGAATGATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTC 180
 QY 181 CTCTTGGCTTTGCTGCTGCTGACCGTTCCAGCTTTCGCTTATGAAGTGGCGCAAGTG 240
 DB |||||
 DB 181 CTCTTGGCTTTGCTGCTGCTGACCGTTTCAGCTTTCGCTTATGAAGTGGCGCAAGTG 240
 QY 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 DB |||||
 DB 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 QY 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 DB |||||
 DB 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 QY 361 TGCTGGGTAGCGCTCACCCCGCGCTGCGAGTGGAGAACGCCAGCGTCCCGACGACA 420
 DB |||||
 DB 361 TGCTGGGTAGCGCTCACCCCGCGCTGCGAGTGGAGAACGCCAGCGTCCCGACGACA 420
 QY 421 ATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 DB |||||
 DB 421 ATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCGCACATACCGGGTCCACCGTATGGGAT 540
 DB |||||
 DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCGCACATACCGGGTCCACCGTATGGGAT 540
 QY 541 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 599
 DB |||||
 DB 541 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 599

RESULT 6
 AAL48928
 ID AAL48928 standard; DNA; 636 BP.

XX AAL48928;
 AC AC
 XX AC
 DT 24-OCT-2002 (first entry)
 XX

DE Hepatitis C virus clone HCC140 E1 protein coding sequence.
 XX
 XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 XX
 OS Hepatitis C virus.
 XX
 XX WO20025548-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EP00219.
 XX

PR 11-JAN-2001; 2001US-260699P.
 PR 30-AUG-2001; 2001US-315768P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Maertens G, Bosman F, Buyse M;
 XX
 XX WPI; 2002-599657/64.
 DR P-PSDB; AAO18669.
 DR
 XX New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection
 XX
 XX Example 2; Page 179-180; 243pp; English.
 PS
 XX The present invention relates to new therapeutic vaccine compositions for
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.

XX SQ Sequence 636 BP; 119 A; 203 C; 174 G; 140 T; 0 other;

Query Match 98.6%; Score 597.4; DB 24; Length 636;
 Best Local Similarity 99.8%; Pred. No. 2.4e-150;
 Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTGATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGGTACATT 60
 DB |||||
 DB 1 ATGTTGGGTAAGGTGATCGATACCTTACATGCGGCTTCGCCACCTCGTGGGGTACATT 60
 QY 61 CGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGCGCATGCGCGTTCGGGTT 120
 DB |||||
 DB 61 CGCTCGTGGGCGCCCTAGGGGGCGTGCAGGGCCCTGCGCATGCGCGTTCGGGTT 120
 QY 121 CTGGAGGACGGCGTGAATGATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTC 180
 DB |||||
 DB 121 CTGGAGGACGGCGTGAATGATGCAACAGGGAATTTGCCGGTTGCTTTCTCTATCTTC 180
 QY 181 CTCTTGGCTTTGCTGCTGCTGACCGTTTCAGCTTTCGCTTATGAAGTGGCGCAAGTG 240
 DB |||||
 DB 181 CTCTTGGCTTTGCTGCTGCTGACCGTTTCAGCTTTCGCTTATGAAGTGGCGCAAGTG 240
 QY 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 DB |||||
 DB 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 QY 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 DB |||||
 DB 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 QY 361 TGCTGGGTAGCGCTCACCCCGCGCTGCGAGTGGAGAACGCCAGCGTCCCGACGACA 420
 DB |||||
 DB 361 TGCTGGGTAGCGCTCACCCCGCGCTGCGAGTGGAGAACGCCAGCGTCCCGACGACA 420
 QY 421 ATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 DB |||||
 DB 421 ATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCGCACATACCGGGTCCACCGTATGGGAT 540
 DB |||||
 DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCGCACATACCGGGTCCACCGTATGGGAT 540
 QY 541 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 599
 DB |||||
 DB 541 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 599

CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 795 BP, 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 85.2%; Score 516.2; DB 24; Length 795;
Best Local Similarity 88.9%; Pred. No. 1.5e-128;
Matches 600; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 1 ATCTGGGTAAGGTCTATCGATACCTTACATCGGCTTCGCGACCTCGTGGGGTACATT 60
DB 1 ATCTGGGTAAGGTCTATCGATACCTTACATCGGCTTCGCGACCTCGTGGGGTACATT 60
QY 61 CCCTGCTCGGCGCCCTCCCTAGGGGCGCTGCGAGGCGCTGCGCATGGCGGCTT 120
DB 61 CCGCTCGTGGCGCCCTCCCTAGGGGCGCTGCGAGGCGCTGCGCATGGCGGCTT 120
QY 121 CTGGAGGACGGCGTGAACATATGCAAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAACATATGCAAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG 240
QY 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGG 300
DB 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGG 300
QY 301 GACATGATCATGACACCGCGGCTGCGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACCGCGGCTGCGCTGCGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGCTCGAGCTAGGAAACGCGAGCTGCCACACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGCTCGAGCTAGGAAACGCGAGCTGCCACACGACA 420
QY 421 ATACGACGCGAGCTCGAT-----TCCAGACTGTTTCCATCTCGCTCGCTCGCGG 468
DB 421 ATACGACGCGAGCTCGAT-----TCCAGACTGTTTCCATCTCGCTCGCTCGCGG 468
QY 439 -----TCCAGACTGTTTCCATCTCGCTCGCTCGCTCGCGG 540
DB 439 -----TCCAGACTGTTTCCATCTCGCTCGCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGT 528
DB 469 CATGAGACGGTGCAGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGT 528
QY 541 CATGAGACGGTGCAGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGT 600
DB 541 CATGAGACGGTGCAGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACGT 600
QY 529 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 588
DB 529 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 588
QY 601 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 660
DB 601 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 660
QY 589 CTCGGATCTCTTAA 603
DB 589 CTCGGATCTCTTAA 603
QY 661 CTCGGATCTCTTAA 675
DB 661 CTCGGATCTCTTAA 675

RESULT 11

AAAL48939
ID AAL48939 standard; DNA; 2082 BP.

XX AC AAL48939;

XX DT 24-OCT-2002 (first entry)

XX DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.

XX KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;

XX KW virucide; immunostimulant; vaccine; ds.

XX OS Hepatitis C virus.

XX PN WO20025548-A2.

DB 361 TCGTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGACGGTCCCAACACGACA 420
QY 421 ATACGACGCGACGCTCGAT-----TCCAGACTGTTTCCATCTCGCTCGCTCGCGG 468
DB 421 ATACGACGCGACGCTCGATTTGCTCGTGGGGCGGCTGCTTTCTGTTCGCTATGTACGTG 480
QY 439 -----TCCAGACTGTTTCCATCTCGCTCGCTCGCTCGCGG 468
DB 439 -----TCCAGACTGTTTCCATCTCGCTCGCTCGCTCGCGG 468
QY 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACGT 528
DB 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACGT 528
QY 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACGT 600
DB 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACGGGTCAACGT 600
QY 529 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 588
DB 529 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 588
QY 601 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 660
DB 601 ATGGCTTGGGATATGATGATGAACCTGCTGCGCTACAAAGGCGCTCGTGTATCGAGCTG 660
QY 589 CTCGGATCTCTTAA 603
DB 589 CTCGGATCTCTTAA 603
QY 661 CTCGGATCTCTTAA 675
DB 661 CTCGGATCTCTTAA 675

RESULT 10

AAAL48914
ID AAL48914 standard; DNA; 795 BP.

XX AC AAL48914;

XX DT 24-OCT-2002 (first entry)

XX DE Hepatitis C virus clone HCC110A E1 protein coding sequence.

XX KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;

XX KW virucide; immunostimulant; vaccine; ds.

XX OS Hepatitis C virus.

XX PN WO20025548-A2.

XX PD 18-JUL-2002.

XX PF 11-JAN-2002; 2002WO-EP00219.

XX PR 11-JAN-2001; 2001US-260699P.

XX PR 30-AUG-2001; 2001US-315768P.

XX PA (INNO-) INNOGENETICS NV.

XX PI Maertens G, Bosman F, Buyse M;

XX DR WPI; 2002-599657/64.

XX DR P-PSDB; AAO18661.

XX New therapeutic vaccine compositions comprising at least one purified recombinant hepatitis C virus (HCV) single or specific oligomeric recombinant envelope protein E1 or E2, useful for immunizing humans from HCV infection.

XX Example 2; Page 161-162; 243pp; English.

XX The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunizing humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence

XX PD 18-JUL-2002.
 XX PF 11-JAN-2002; 2002WO-EP00219.
 XX PR 11-JAN-2001; 2001US-260699P.
 XX PR 30-AUG-2001; 2001US-315768P.
 XX PA (INNO-) INNOGENETICS NV.
 XX PI Maertens G, Bosman F, Buyse M;
 DR WPI; 2002-599657/64.
 DR P-PSDB; AA018678.
 XX PT New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans
 PT from HCV infection
 XX PS Example 2; Page 206-209; 243pp; English.
 XX The present invention relates to new therapeutic vaccine compositions for
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope protein selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention.
 XX
 XX SQ Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

Query Match 84.7%; Score 513.2; DB 24; Length 2082;
 Best Local Similarity 88.8%; Pred. No. 1.3e-127;
 Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 4 TTGGGTAAGTCAATGATACCCCTTACATCGCGGCTTCGCGACCTCGTGGGTACATTCGG 63
 DB 4 TTGGGTAAGTCAATGATACCCCTTACATCGCGGCTTCGCGACCTCGTGGGTACATTCGG 63
 QY 64 CTGTCGCGCGCCCCCTAGCGGGCGCTGCGAGGGCCCTGCGCATGCGCGTCCGGTTCTG 123
 DB 64 CTGTCGCGCGCCCCCTAGCGGGCGCTGCGAGGGCCCTGCGCATGCGCGTCCGGTTCTG 123
 QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTCCTC 183
 DB 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTCCTC 183
 QY 184 TTGGCTTGTCTGCTGCTGACCGTTCAGCTTCGCTATGAAGTGGCAGCGTTC 243
 DB 184 TTGGCTTGTCTGCTGCTGACCGTTCAGCTTCGCTATGAAGTGGCAGCGTTC 243
 QY 244 GGGATGATACATGCTACAGAACGACTGCTCCAACTCAAGCATTCGTATGAGGAGCGGAC 303
 DB 244 GGGATGATACATGCTACAGAACGACTGCTCCAACTCAAGCATTCGTATGAGGAGCGGAC 303
 QY 304 ATGATCATGCACACCCCGGGTGCCTGCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
 DB 304 ATGATCATGCACACCCCGGGTGCCTGCGCTTCGGGAGAACAACTCTTCCCGCTGC 363
 QY 364 TGGGTAGCGCTACCCCGAGCTCGAGTAGGAAAGCCAGCGTCCCGACACGACAATA 423
 DB 364 TGGGTAGCGCTACCCCGAGCTCGAGTAGGAAAGCCAGCGTCCCGACACGACAATA 423
 QY 424 CGACGCCACGTCGAT----- 438
 DB 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGGG 483

QY 439 -----TCCAGCTGTTTCCACATCTCGCTCCCGGCAT 471
 DB 484 GACCTCTCGGATCTGTCTTCTCTCGCTCCCGGCAT 543
 QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGGTATG 531
 DB 544 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATAACGGGTACCGGTATG 603
 QY 532 GCTTGGGATATGATGATGAATGAACTGGTGCCTCAACAGCGCCCTGGTGTATGCGAGCTGCTC 591
 DB 604 GCTTGGGATATGATGATGAATGAACTGGTGCCTCAACAGCGCCCTGGTGTATGCGAGCTGCTC 663
 QY 592 CGGATCCTCTAA 603
 DB 664 CGGATCCCAAA 675
 RESULT 12
 AAT12973
 ID AAT12973 standard; DNA; 2086 BP.
 XX
 AC AAT12973;
 XX
 DT 24-SEP-1996 (first entry)
 XX
 DE HCV E1 construct HCC165.
 XX
 KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9604385-A2.
 XX
 PD 15-FEB-1996.
 XX
 PF 31-JUL-1995; 95WO-EP03031.
 XX
 PR 29-JUL-1994; 94EP-0870132.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Bosman F, Buyse M, De Martynoff G, Maertens G;
 DR WPI; 1996-129401/13.
 XX
 PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of di: sulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 PS Claim 23; Fig 21; 146pp; English.
 XX
 CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the
 CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.
 XX

15-FEB-1996.

31-JUL-1995; 95WO-EP03031.

29-JUL-1994; 94EP-0870132.

(INNO-) INNOGENETICS NV.

Bosman F, Buyse M, De Martynoff G, Maertens G; WPI; 1996-129401/13.

Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of di: sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

Claim 23; Fig 21; 146pp; English.

AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;

Query Match 84.7%; Score 513.2; DB 17; Length 2433;
Best Local Similarity 88.8%; Pred. No. 1.3e-127;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;

QY 4 TTGGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCCACCTCGTGGGTACATTCG 63
Db 355 TTGGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCCACCTCGTGGGTACATTCG 414
QY 64 CTCGTGGGCGCCCTAGGGGGCGCTGCCAGGCGCTTCGCCACCTCGTGGGTACATTCG 123
Db 415 CTCGTGGGCGCCCTAGGGGGCGCTGCCAGGCGCTTCGCCACCTCGTGGGTACATTCG 474
QY 124 GAGACGGCGGTGAACCTATGCAACAGGGAATTTGCGCGGTCTCTTCTATCTTCCTC 183
Db 475 GAGACGGCGGTGAACCTATGCAACAGGGAATTTGCGCGGTCTCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGTCTGTCTGACCGCTTCAGCTTCGCTTATGAAGTGCACAGTGTCC 243
Db 535 TTGGCTTTGCTGTCTGTCTGACCGCTTCAGCTTCGCTTATGAAGTGCACAGTGTCC 594
QY 244 GGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCGGAC 303
Db 595 GGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCGGAC 654
QY 304 ATGATCATGACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGTTCG 363
Db 655 ATGATCATGACACCCCGGGTGGCTGCGCTTCGGGAGAACAACTCTTCCCGTTCG 714
QY 364 TGGGTAGCGGTCAACCCCGGCGTGCAGTGAAGACCGCGCTCCACACGACCAATA 423
Db 715 TGGGTAGCGGTCAACCCCGGCGTGCAGTGAAGACCGCGCTCCACACGACCAATA 774
QY 424 CGACGCCACGTGCGAT----- 438
QY 424 CGACGCCACGTGCGAT----- 438
QY 439 -----TCCAGCTGTTCACCATCTCGCTCGCGGAT 471
Db 484 GACCTCTCGGATCTGTCTTCTCGCTCCAGCTGTTCACCATCTCGCTCGCGGAT 543
QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 531
Db 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATG 603
QY 532 GCTTGGGATATGATGATGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
Db 604 GCTTGGGATATGATGATGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY 592 CGGATCTCTAA 603
Db 664 CGGATCTCTAA 675
RESULT 13
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX AC AAT12974;
XX DT 25-SEP-1996 (first entry)
XX DE HCV E1 construct HCC166.
XX KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX KW ss.
XX OS Hepatitis C virus.
XX PN W09604385-A2.
XX

Db 775 CGAGCCACGTCGATTTGCTGTTGGGCGGCTGCTTTCTTCCGCTATGTACGTGGG 834
QY 439 -----TCCAGCTGTTACCATCTCGCCCTCGCCGGCAT 471
Db 835 GACCTCTGGGATCTGTCCTCTCCAGCTGTTACCATCTCGCCCTCGCCGGCAT 894
QY 472 GAGACGGTGAGGACTGCAATTGCTCAATATCCCGGCCACATACAGGGTCAACGATG 531
Db 895 GAGACGGTGAGGACTGCAATTGCTCAATATCCCGGCCACATACAGGGTCAACGATG 954
QY 532 GCTTGGGATATGATGTAAGTCTGCTCGCTTACAGCGCCCTGCTGATGCGAGCTGCTC 591
Db 955 GCTTGGGATATGATGTAAGTCTGCTCGCTTACAGCGCCCTGCTGATGCGAGCTGCTC 1014
QY 592 CGGATCTCTTAA 603
Db 1015 CGGATCCCA 1026

RESULT 14
AAL48940
ID AAL48940 standard; DNA; 2434 BP.
XX
AC AAL48940;
XX
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
XX
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus..

XX
XX WO200255548-A2.
XX
XX 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
XX
XX 30-AUG-2001; 2001US-315766P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buysse M;
XX
XX WPI; 2002-599657/64.
XX
XX P-PSDB; AAO18679.
XX
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection

XX
XX Example 2; Page 212-215; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX in vitro monitoring of HCV disease or prognosing the response to
XX treatment of patients suffering from HCV infection. The present sequence
XX is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;

Query Match 82.9%; Score 502.2; DB 24; Length 2434;
Best Local Similarity 88.7%; Pred. No. 1.2e-124;
Matches 597; Conservative 0; Mismatches 3; Indels 73; Gaps 2;
QY 4 TTGGGTAAGGTGTCATCGATACCCCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 63
Db 355 TTGGGTAAGGTGTCATCGATACCCCTTACATGCGGCTTCGCGACCTCGTGGGTACATTCGG 414
QY 64 CTGCTCGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGCGCATGCGCGTCCGGGTCTG 123
Db 415 CTGCTCGGCGCCCTAGGGGCGCTGCGAGGCGCTGCGCGCATGCGCGTCCGGGTCTG 474
QY 124 GAGGACGGCGTGAACTATGATGCAACAGGGAATTTCCCGGTTGCTCTTCTATCTTCCTC 183
Db 475 GAGGACGGCGTGAACTATGATGCAACAGGGAATTTCCCGGTTGCTCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGTCCTG-TCTGACCGTTCAGCTTCGCGTTCGCGTATGAAAGTGGCAACGTGTC 242
Db 535 TTGGCTTTGCTGTCCTG-TCTGACCGTTCAGCTTCGCGTTCGCGTATGAAAGTGGCAACGTGTC 594
QY 243 CGGGATGTACCATGTTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGA 302
Db 595 CGGGATGTACCATGTTCACGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGA 654
QY 303 CATGATCATGACACACCCCGGGTGGTGGCTGCGGTTCGGGAGAAACAATCTTTCCGCTG 362
Db 655 CATGATCATGACACACCCCGGGTGGTGGCTGCGGTTCGGGAGAAACAATCTTTCCGCTG 714
QY 363 CTGGGTAGCGCTCACCCCGCGCTGCGAGCTAGGACGCGGTCGCCACCGACGACAT 422
Db 715 CTGGGTAGCGCTCACCCCGCGCTGCGAGCTAGGACGCGGTCGCCACCGACGACAT 774
QY 423 ACAGCCGACGCTCGAT-----TTCGGGGGGGCTGCTTTCTGTCGCTATGTACGTGGG 438
Db 775 ACAGCCGACGCTCGATTTGCTTTCGGGGGGGCTGCTTTCTGTCGCTATGTACGTGGG 834
QY 439 -----TCCAGCTGTTTACCATCTCGCCCTCGCCGGCA 470
Db 835 GGACCTCTCGGATCTGCTTCTTCCTCGCTCCAGCTGTTCCACCATCTCGCCGGCA 894
QY 471 TGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATACCGGTACCGTAT 530
Db 895 TGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATACCGGTACCGTAT 954
QY 531 GGCTTGGGATATGATGATGAATGCTCGCTTACAGCGCCCTGCTGCTATCGCAGCTGCT 590
Db 955 GGCTTGGGATATGATGATGAATGCTCGCTTACAGCGCCCTGCTGCTATCGCAGCTGCT 1014
QY 591 CCGGATCTCTTAA 603
Db 1015 CCGGATCCCA 1027

RESULT 15
AAT12706
ID AAT12706 standard; DNA; 633 BP.
XX
XX AC AAT12706;
XX
XX DT 23-SEP-1996 (first entry)
XX
XX DE HCV E1 construct HCC111A.
XX
XX KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX
XX OS Hepatitis C virus.
XX
XX PN WO9604385-A2.
XX
XX PD 15-FEB-1996.
XX

Mon Dec 22 13:28:41 2003

QY 439 -----TCCAGCTGTTCCACCATCTCGCTCGCGG 468
Db 481 GGGGATCTCTGGGATCTCTCTTCCTCGTCTCCAGCTGTTCCACCATCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATACGGGTACCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCACATACGGGTACCGT 600
QY 529 ATGGCTTGGGATATGATGAAGTGGT 556
Db 601 ATGGCTTGGGATATGATGAAGTGGT 628
Search completed: December 19, 2003, 18:51:18
Job time : 170.534 secs

PF 31-JUL-1995; 95WO-BP03031.
XX PR
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di:sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
XX Claim 23; Fig 21; 146pp; English.
XX
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;
SQ
Query Match 75.3%; Score 456.4; DB 17; Length 633;
Best Local Similarity 86.8%; Pred. No. 1.5e-112;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCGTGGGTACATT 60
Db 1 ATGTTGGGTAGGTATCATGATACCTTACATCGCGCTTCGCGGACCTCATGGGTACATT 60
QY 61 CCGCTCGTCGCGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGGCGATGGCGTCCGGTT 120
Db 61 CCGCTCGTCGCGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGGGCGATGGCGTCCGGTT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGTCTGCTCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTGGCTTTGCTGTCTGCTCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTACGAAACGACTGCTCAACTCAAGATTTGTGTATGAGCAGCG 300
Db 241 TCCGGGATGTACCATGTACGAAACGACTGCTCAACTCAAGATTTGTGTATGAGCAGCG 300
QY 301 GACATGATCATGACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGACACCCCGGGTGGTGGCTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTACCCCGAGCTCGCAGCTAGGAAACGCGTCCCGACGACA 420
Db 361 TGCTGGGTAGCGCTACCCCGAGCTCGCAGCTAGGAAACGCGTCCCGACGACA 420
QY 421 ATACGACGCCACGTCGAT----- 438
Db 421 ATACGACGCCACGTCGATTTTCTGTTGGGGCGGCTGCTTCTTCTGTCGCTATGACGTG 480

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1558.97 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title:

US-09-899-303A-25

Perfect score: 606

Sequence: 1 ATGTTGGTAGGTCATCGA.....TGCTCCGGATCCTCTATAG 606

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_estbha:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rnd:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gsl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	68.2	11.3	488	9 AV755731	AV755731
C 2	54.6	9.0	492	9 AV758366	AV758366
C 3	41.6	6.9	502	12 BI879124	BI879124 fm04e08.y
C 4	40.6	6.7	275	9 AV835132	AV835132

5	40.6	6.7	402	9	AV392783	AV392783
6	40.6	6.7	551	9	AV392165	AV392165
7	40.6	6.7	552	12	BI996341	BI996341
8	40.6	6.7	584	12	BI727879	BI727879
9	40.4	6.7	1201	13	BI356664	BI356664
10	40.2	6.6	1162	12	BM918259	BM918259
11	40	6.6	1201	9	AL513886	AL513886
12	39	6.4	359	12	BU252669	BU252669
13	39	6.4	375	12	BU246716	BU246716
14	39	6.4	840	29	CC335916	CC335916
15	39	6.4	873	14	CD446071	CD446071
16	38.6	6.4	925	29	CNS0091P	AL053013
17	38.4	6.3	636	12	BI960110	BI960110
18	38.4	6.3	702	14	CD432549	CD432549
19	38.4	6.3	970	29	CNS010C9	CNS010C9
20	38.4	6.3	987	29	CNS015VX	CNS015VX
21	38.2	6.3	533	6	AU192776	AU192776
22	38.2	6.3	538	6	AU193705	AU193705
23	38.2	6.3	544	6	AU190971	AU190971
24	38.2	6.3	544	6	AU192419	AU192419
25	38.2	6.3	1270	12	BG968359	BG968359
26	38	6.3	354	14	CB966525	CB966525
27	38	6.3	1201	13	BI381961	BI381961
28	37.8	6.2	435	14	C72860	C72860
29	37.8	6.2	533	29	CC010084	CC010084
30	37.8	6.2	659	29	CC405164	CC405164
31	37.8	6.2	826	29	BZ736582	BZ736582
32	37.8	6.2	895	29	CC359028	CC359028
33	37.8	6.2	925	29	CC359026	CC359026
34	37.8	6.2	940	29	CC010085	CC010085
35	37.8	6.2	951	29	CC405167	CC405167
36	37.6	6.2	431	9	AV639153	AV639153
37	37.4	6.2	360	9	AJ473805	AJ473805
38	37.4	6.2	637	13	BQ293470	BQ293470
39	37.4	6.2	641	13	BQ172543	BQ172543
40	37.4	6.2	650	14	CA828039	CA828039
41	37.4	6.2	834	29	BZ641450	BZ641450
42	37.4	6.2	841	29	BZ641457	BZ641457
43	37.4	6.2	856	29	BZ578381	BZ578381
44	37.4	6.2	872	29	BZ550111	BZ550111
45	37.2	6.1	563	10	BE490055	BE490055

ALIGNMENTS

RESULT 1	AV755731/c	488 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	cDNA clone	BMFAK03 5', mRNA sequence.	
DEFINITION	AV755731				
ACCESSION	AV755731				
VERSION	AV755731.1	GI:10913579			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 488)				
AUTHORS	Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.				
TITLE	Homo sapiens cDNA BM clones				
JOURNAL	Unpublished				
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers				

```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
116 a 134 c 137 g 97 t 4 others

BASE COUNT
ORIGIN
Query Match 11.3%; Score 68.2; DB 9; Length 488;
Best Local Similarity 72.3%; Pred. No. 3.6e-07;
Matches 102; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 445 CTGTTCCACATCTCGCTCGCGGCATGAGACGGTGAGAGCTCAATGTCGAATCTAT 504
DB 403 CAGCTGATCATCTGGGCTCAGCACCATGAGTTGTGCGATGATGCACTGCTCATCTAT 344

QY 505 CCCGGCCACATACCGGGTCACCGTATG-GCTTGGATATGATGATGAACCTGGTCGCTAC 563
DB 343 CCTGGCGCCATCACTGCACACCGTATGACATGGGACATGATGATGAACCTGGTCGCTAC 284

QY 564 AACGGCCCTGGTGGTATCGCA 584
DB 283 CGTGTATGATCATGGCGTA 263

RESULT 2
AV758366/c
LOCUS AV758366 BM Homo sapiens CDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens CDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
124 a 128 c 125 g 112 t 3 others

BASE COUNT
ORIGIN
Query Match 9.0%; Score 54.6; DB 9; Length 492;
Best Local Similarity 67.2%; Pred. No. 0.0012;
Matches 92; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

```



```

QY 342 GGAGAACAACTCTCCCGTCTGGGT 368
Db 410 CAAGAAGACGAAGGCGAGCTACATGAT 436

RESULT 8
BI727879
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 584)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
source
1..584
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
BASE COUNT 106 a 188 c 197 g 93 t
ORIGIN
Query Match 6.7%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. No. 5.3;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGGTACATTCGCTCGCGGCCCCCTAGGGGGCGTGCAGGCGCCT 101
Db 47 CGAGTCTATCTGTTGCGGGGACTGCCAACATGAAGAGCGTCTCAGCGAAT 106

QY 102 GCGCATGCGCTCGGGTTCGAGAGACGCGTGAACACTGCAACAGGAAATTCGCCGG 161
Db 107 GCGCGGCGCGCGAGTGGAGGCGCGCTACGCGACGAGTTCGTGAGCTTGGGCG 166

QY 162 TTGCTCTTCTATCTCTCTGCTTGGCTTGTGCTCTGACCGTTCCAGCTTCGCG 221
Db 167 CGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCAAGACCCAG 226

QY 222 TTATGAGTGGCAACGTCGTCGGGATGTACCATGTACGAACGACTGTCTCAACTCAAG 281
Db 227 CTTGCGCGTTCGCTGCGTGGGCACTCGCTGGCGGCGGCACCGCGGCTGCTGCGAT 286

QY 282 CATTGTGTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGTGGCTGCTGCTG 341
Db 287 CCTGATGACCAACGAGGAGGATTTGCGCGCGCATCTACGCGGCGTGCCTCATGCCGG 346

QY 342 GGAGAACAACTCTTCCCGTCTGGGT 368
Db 347 CAAGAAGACGAAGGCGAGCTACATGAT 373

RESULT 9
BI727879
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Chlamydomonas reinhardtii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1015CA02NP1.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YB03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 6.7%; Score 40.4; DB 13; Length 1201;
Best Local Similarity 10.4%; Pred. No. 7.4;
Matches 52; Conservative 233; Mismatches 212; Indels 3; Gaps 1;

QY 34 GGCCTTCGCGACCTCGTGGGTACATTCGCTCGCGGCCCCCTAGGGGGCGCTGCC 93
Db 618 GNTNTSSSSSTNNNNSSSSNNNTTBTBTSSSSSTSSSTSSSTSSSTSSST 677

QY 94 AGGCGCTCGCGGCGTCCGGTCTCGGAGGACGCGTGAACATGATCAACAGGAAT 153
Db 678 SSSSSSBTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 737

QY 154 TTGCCCGGTGCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 213
Db 738 TTKSSSSSTBSSTSTBTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 797

QY 214 GCTTCGCTTATGAAGTGGCAACGTCGCGGAGTGATCACCATGTACGACGACGCTCC 273
Db 798 TBSMTSSSBTCTSSSSSSSSSTTSTSTSTSTSTSTSTSTSTSTSTSTSTST 854

```

Mon Dec 22 13:28:43 2003

```

QY 274 AACTCAAGCATTGTATGAGCGAGCGACATGATCATGCACACCCCGGGTGGTGCCC 333
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
855 TSSNTNTTSSSBSSTBSTSTBSWSBBTBTBTBTBTBTBTBTBTBTBTBTBTBTBTBT 914
QY 334 TCGGTTCCGGAGAACAACTCTCCCGCTGCTGGGTAGCGCTCACCCACCGCTGCGAGCT 393
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
915 SSSSTTBSBSSTBSSTBSSTBSSTBSSTBSSTBSSTBSSTBSSTBSSTBSSTBSST 974
QY 394 AGGAACGCCAGCGTCCCGACACGACAAATAGACCCAGCTCGATTCCAGCTTTCACC 453
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
975 STTBSSTTTTSSSTTSSSTTSSSTTSSSTTSSSTTSSSTTSSSTTSSSTTSSST 1034
QY 454 ATCTCGCTCCCGCATGAGCGGTGAGGACTGCAATGTCTCAATCTATCCCGGCCAC 513
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1035 SNKSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSST 1094
QY 514 ATAACGGGTCCACCGTATGCG 533
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1095 SBTSSSTTTTSSSATBSB 1114

RESULT 10
LOCUS BM918259 1162 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649
5', mRNA sequence.
ACCESSION BM918259
KEYWORDS EST.
SOURCE BM918259.1 GI:19368638
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1162)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcgabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1CM2016 row: n column: 18
High quality sequence stop: 567.
FEATURES
Location/Qualifiers
1..1162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5485649"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pORF7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT 224 a 499 c 240 g 198 t 1 others
ORIGIN
Query Match 6.6%; Score 40.2; DB 12; Length 1162;
Best Local Similarity 54.4%; Pred. No. 8.2;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 313 CACACCCCGGGTGGTGCCCTCGGTTCGGGAGACAACTCTCCCGCTGCTGGGTAGCG 372

```

```

Db 715 CCCCCCCCCGGGTCTCCTCCACACCCCGGACCCCAAAACCCCGGACCGCTCC 774
QY 373 CTCACCCCCACACCTCGCAGCTAGGAACCGCAGCGTCCACACGACAAATAGGAGCCAC 432
Db 775 CCCATTCGCCACGACCGACCCCGCCCATATCCGCCCCCTACCGATCACCCTACCCAC 834
QY 433 GTTCGATTCACGCTGTTTACCATCTCGCC 461
Db 835 GCTGTATCCCGCCCTCGCACCAACCCCGCC 863

RESULT 11
LOCUS AL513886/c 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION AL513886 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA006ZG08
5-PRIME, mRNA sequence.
ACCESSION AL513886
VERSION AL513886.2 GI:30463771
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced gi:12777380.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4924.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA006ZG08RPI&cluster=4924.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBA006ZG08RPI.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA006ZG08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 201 a 311 c 349 g 146 t 194 others
ORIGIN
Query Match 6.6%; Score 40; DB 9; Length 1201;
Best Local Similarity 26.5%; Pred. No. 9.4;
Matches 103; Conservative 104; Mismatches 179; Indels 3; Gaps 1;
QY 32 CGCGCTTCCCGACCTCGTGGGTACATTCCTCGTGGCGCCCCCTAGGGGGCGTG 91
Db 1129 GSCGCGTGCBSBSCNTYKKGKBSSSSSCCSVSSSGSGSCSSSCCGGGGGGGGGG 1070
QY 92 CCAGGGCCCTCGCGCATGCGCTCCGGTCTCTGGAGACCGCTGAACATGCAACAGGA 151
Db 1069 CCSGKGKGGGKSSSSSGSGCGCCCGGG---GGSSSGSGSGGSCCWAASASAYKKGK 1013
QY 152 ATTTCGCCGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 211
Db 1012 GKKTITTTTWWMAATHTTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 953

```

```

QY 212 CAGCTTCGGCTTATGAGTGGCGCAACGCTGTCGGGATGTACCATGTGCACAGCACTGCT 271
Db 952 GACSCSCCGCAADAACGAGMDGAMGKSTGTGSCCTTSRCRWGGTTTSGMMVGCATTY 893
QY 272 CCAACTCAAGCATTTGTATGAGCGAGCATATGATCATGCACACACCCCGGCTGGTCC 331
Db 892 AYBSYTGTRRWTWTGTSBTCTYASGSGMYSSKKBKCKCMAYAACSCAGASCST 833
QY 332 CTGCGTTCCGGAGAACAACTTCTCCCGCTGCTGGTAGCGTCAACCCACCGCTGCAG 391
Db 832 SGCSRGKTKTTTGTGCTGTTGAAGASMBRTWAGGGGGGGCCCGCCYCSMCCGCCCVB 773
QY 392 CTAGGAACGCCAGCGTCCCCACACGACA 420
Db 772 BBCCWCHCTKCKMCCRGACTYCCCA 744

RESULT 12
BJ252669/c
LOCUS
DEFINITION
  BJ252669 Y. Ogihara unpublished cDNA library, Wh_f Triticum
  aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION
  BJ252669
VERSION
  BJ252669.1 GI:20061830
KEYWORDS
  EST.
SOURCE
  Triticum aestivum (bread wheat)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Triticum.
  1 (bases 1 to 359)
REFERENCE
  Ogihara, Y. and Murai, K.
  Expressed genes in Triticum aestivum
  Unpublished
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
  Location/Qualifiers
    1..359
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="Chinese Spring"
    /db_xref="taxon:4565"
    /clone="whf25g19"
    /tissue_type="spike at flowering date"
    /dev_stages="Feekes' scale 10.5.1"
    /clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
    70 a 115'c 107 g 67 t

BASE COUNT
  70 a 115'c 107 g 67 t
ORIGIN
  70 a 115'c 107 g 67 t

FEATURES
  source
  Query Match 6.4%; Score 39; DB 12; Length 359;
  Best Local Similarity 58.0%; Pred. No. 12;
  Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTGGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTTAGG 83
Db 297 CTTCAAGTGCACAGCGCGTCTGGAAGCGCTCAGGGCGGTGACGCGCTGCGCGTCGG 238
QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTTGAGGAGCGGCGTCAACTATG 142
Db 237 GGACGCCGCGACCCCTGGCGCAGGACGTGCACGTGCTGCCGTGCACGTGCCCAAG 179

Query Match 6.4%; Score 39; DB 12; Length 359;
Best Local Similarity 58.0%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTGGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTTAGG 83
Db 297 CTTCAAGTGCACAGCGCGTCTGGAAGCGCTCAGGGCGGTGACGCGCTGCGCGTCGG 238
QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTTGAGGAGCGGCGTCAACTATG 142
Db 237 GGACGCCGCGACCCCTGGCGCAGGACGTGCACGTGCTGCCGTGCACGTGCCCAAG 179

RESULT 13
BJ246716
LOCUS
DEFINITION
  BJ246716 Y. Ogihara unpublished cDNA library, Wh_f Triticum
  aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION
  BJ246716
VERSION
  BJ246716.1 GI:20058228
KEYWORDS
  EST.
SOURCE
  Triticum aestivum (bread wheat)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Triticum.
  1 (bases 1 to 375)
REFERENCE
  Ogihara, Y. and Murai, K.
  Expressed genes in Triticum aestivum
  Unpublished
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
  Location/Qualifiers
    1..375
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="Chinese Spring"
    /db_xref="taxon:4565"
    /clone="whf25g19"
    /tissue_type="spike at flowering date"
    /dev_stages="Feekes' scale 10.5.1"
    /clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
    81 a 107 c 110 g 77 t

BASE COUNT
  81 a 107 c 110 g 77 t
ORIGIN
  81 a 107 c 110 g 77 t

FEATURES
  source
  Query Match 6.4%; Score 39; DB 12; Length 375;
  Best Local Similarity 58.0%; Pred. No. 12;
  Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTGGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTTAGG 83
Db 36 CTTCAAGTGCACAGCGCGCTCTGGAAGCGCTCAGGGCGGTGACGCGCTGCGCGTCGG 95
QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTTGAGGAGCGGCGTCAACTATG 142
Db 96 GGACGCCGCGACCCCTGGCGCAGGACGTGCACGTGCTGCCGTGCACGTGCCCAAG 154

RESULT 14
CC335916
LOCUS
DEFINITION
  OGUAJ60TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0393124,
  genomic survey sequence.
ACCESSION
  CC335916
VERSION
  CC335916.1 GI:30805329
KEYWORDS
  GSS.
SOURCE
  Zea mays
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 840)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
  , A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
  , R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
  Location/Qualifiers
    1..840

FEATURES
  source
  Query Match 6.4%; Score 39; DB 12; Length 375;
  Best Local Similarity 58.0%; Pred. No. 12;
  Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTGGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTTAGG 83
Db 36 CTTCAAGTGCACAGCGCGCTCTGGAAGCGCTCAGGGCGGTGACGCGCTGCGCGTCGG 95
QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTTGAGGAGCGGCGTCAACTATG 142
Db 96 GGACGCCGCGACCCCTGGCGCAGGACGTGCACGTGCTGCCGTGCACGTGCCCAAG 154

RESULT 14
CC335916
LOCUS
DEFINITION
  OGUAJ60TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0393124,
  genomic survey sequence.
ACCESSION
  CC335916
VERSION
  CC335916.1 GI:30805329
KEYWORDS
  GSS.
SOURCE
  Zea mays
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 840)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick
  , A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
  , R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
  Location/Qualifiers
    1..840

FEATURES
  source
  Query Match 6.4%; Score 39; DB 12; Length 375;
  Best Local Similarity 58.0%; Pred. No. 12;
  Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 CCTTACATCGCGCTTGGCGACCTCGTGGGTACATTCCGCTCGTGGCGCCCTTAGG 83
Db 36 CTTCAAGTGCACAGCGCGCTCTGGAAGCGCTCAGGGCGGTGACGCGCTGCGCGTCGG 95
QY 84 GGGCGCTGCAGGCGCCCTGGCGCATGGCTCGGGTTCTTGAGGAGCGGCGTCAACTATG 142
Db 96 GGACGCCGCGACCCCTGGCGCAGGACGTGCACGTGCTGCCGTGCACGTGCCCAAG 154

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 43.7045 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-25

Perfect score: 606

Sequence: 1 ATGTTGGTAAAGTCATCGA.....TGCTCGGATCCTCTAATAG 606

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	606	100.0	606	3	US-08-612-973-25
2	606	100.0	606	3	US-08-927-597-25
3	598.2	98.7	723	3	US-08-612-973-21
4	598.2	98.7	723	3	US-08-927-597-21
5	597.4	98.6	636	3	US-08-612-973-27
6	597.4	98.6	636	3	US-08-927-597-27
7	556	91.7	561	3	US-08-612-973-23
8	556	91.7	561	3	US-08-927-597-23
9	516.2	85.2	795	3	US-08-612-973-5
10	516.2	85.2	795	3	US-08-927-597-5
11	513.2	84.7	2082	3	US-08-612-973-47
12	513.2	84.7	2082	3	US-08-927-597-47
13	513.2	84.7	2433	3	US-08-612-973-49
14	513.2	84.7	2433	3	US-08-927-597-49
15	456.4	75.3	633	3	US-08-612-973-7
16	456.4	75.3	633	3	US-08-927-597-7
17	451.6	74.5	636	3	US-08-612-973-13
18	451.6	74.5	636	3	US-08-927-597-13
19	449.2	74.1	1539	2	US-08-470-426B-17
20	449.2	74.1	1539	2	US-08-470-426B-14
21	444.4	73.3	9595	3	US-09-014-416-4
22	444.4	73.3	9599	3	US-09-014-416-6
23	442.8	73.1	932	1	US-08-081-072-15
24	442.8	73.1	932	1	US-08-449-093A-15
25	442.8	73.1	2116	3	US-08-191-160-21
26	441.2	72.8	9472	4	US-08-150-204B-96
27	441	72.8	1037	1	US-08-462-195-1

28	441	72.8	1037	2	US-08-636-883-1	Sequence 1, Appli
29	441	72.8	1037	3	US-09-127-829-1	Sequence 1, Appli
30	430	71.0	1167	1	US-08-324-977-9	Sequence 9, Appli
31	430	71.0	1167	2	US-08-384-616-9	Sequence 9, Appli
32	430	71.0	1167	2	US-08-904-686A-9	Sequence 9, Appli
33	430	71.0	1167	3	US-09-315-850-9	Sequence 9, Appli
34	430	71.0	1499	1	US-08-324-977-3	Sequence 3, Appli
35	430	71.0	1499	2	US-08-384-616-3	Sequence 3, Appli
36	430	71.0	1499	2	US-08-904-686A-3	Sequence 3, Appli
37	430	71.0	1499	3	US-09-315-850-3	Sequence 3, Appli
38	430	71.0	6039	1	US-08-324-977-11	Sequence 11, Appli
39	430	71.0	6039	2	US-08-384-616-11	Sequence 11, Appli
40	430	71.0	6039	2	US-08-904-686A-11	Sequence 11, Appli
41	430	71.0	9030	3	US-09-315-850-11	Sequence 11, Appli
42	430	71.0	9030	1	US-08-324-977-13	Sequence 13, Appli
43	430	71.0	9030	2	US-08-384-616-13	Sequence 13, Appli
44	430	71.0	9030	2	US-08-904-686A-13	Sequence 13, Appli
45	430	71.0	9030	3	US-09-315-850-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-612-973-25
; Sequence 25, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..603
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..600

```

US-08-612-973-25
Query Match      100.0%; Score 606; DB 3; Length 606;
Best Local Similarity 100.0%; Pred. No. 8.4e-153;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US 08/612,973
APPLICATION NUMBER:
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..603
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..600
US-08-927-597-25

Query Match      100.0%; Score 606; DB 3; Length 606;
Best Local Similarity 100.0%; Pred. No. 8.4e-153;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible


```

QY 541 ATGATGATGAAGTGGTGGCTTACACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCTC 600
DB 541 ATGATGATGAAGTGGTGGCTTACACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCTC 600
QY 601 TAATAG 606
DB 601 TAATAG 606

RESULT 3

US-08-612-973-21
; Sequence 21, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..717
; US-08-612-973-21

Query Match 98.7%; Score 598.2; DB 3; Length 723;
Best Local Similarity 99.5%; Pred. No. 1.1e-150;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTTGGTGAAGTGCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGTGAAGTGCATCGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CCCTCGTGGCGCCCTAGGGGGCTGCCAGGGCCCTGGCGATGCGTCCGGTT 120
DB 61 CCCTCGTGGCGCCCTAGGGGGCTGCCAGGGCCCTGGCGATGCGTCCGGTT 120

QY 121 CTGGAGGAGGGCGTGAACATATGCAACAGGGAATTTCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGAGGGCGTGAACATATGCAACAGGGAATTTCCCGGTTGCTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGAGCCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGTG 240
DB 181 CTCTTGGCTTTGCTGCTGCTGAGCCGTTCCAGCTTCCGCTTATGAAGTGGCGCAAGTG 240
QY 241 TCCGGATGTACCATGTCAAGAACGATGCTTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGATGTACCATGTCAAGAACGATGCTTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATCATGACACACCCCGGTTGGTGGCTTCCGCTTCCGAGAACAACTTCTCCCGC 360
DB 301 GACATCATGACACACCCCGGTTGGTGGCTTCCGCTTCCGAGAACAACTTCTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTGAGAGCTAGGAAAGCCAGCGTCCCAACACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTGAGAGCTAGGAAAGCCAGCGTCCCAACACGACA 420
QY 421 ATACGACGCGCAGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGATGAGCGTG 480
DB 421 ATACGACGCGCAGCTCGATTCCCAAGCTGTTCAACCATCTCGCCTCGCGGATGAGCGTG 480
QY 481 CAGGACTGCAATGCTCAATCTATCCCGGACATTAACCGGTCACCGTATGGTATGGAT 540
DB 481 CAGGACTGCAATGCTCAATCTATCCCGGACATTAACCGGTCACCGTATGGTATGGAT 540
QY 541 ATGATGATGAAGTGGTGGCTTACCAAGCGGCTGCTGATCGCAGCTGCCGATCCTC 600
DB 541 ATGATGATGAAGTGGTGGCTTACCAAGCGGCTGCTGATCGCAGCTGCCGATCCTC 600
QY 601 TAA 603
DB 601 CAA 603

RESULT 4

US-08-927-597-21
; Sequence 21, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:

Mon Dec 22 13:28:42 2003

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..720
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..717
US-08-927-597-21

Query Match 98.7%; Score 598.2; DB 3; Length 723;
Best Local Similarity 99.5%; Pred. No. 1.1e-150;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGSCCTGGCGATGGCGTCCGGGTT 120
DB 61 CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGSCCTGGCGATGGCGTCCGGGTT 120

QY 121 CTGGAGGACGCGTGAACTATGCAACAGGGAATTTGCCGGTTGCTTTCTATCTTC 180
DB 121 CTGGAGGACGCGTGAACTATGCAACAGGGAATTTGCCGGTTGCTTTCTATCTTC 180

QY 181 CTCCTGGCTTTGCTGCTCTGCTGACCGTTCCAGTTCCCGTTATGAAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTGCTGCTCTGCTGACCGTTCCAGTTCCCGTTATGAAGTGGCAACGTG 240

QY 241 TCCGGGATGATACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300
DB 241 TCCGGGATGATACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300

QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAACACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAACACTCTTCCCGC 360

QY 361 TCGTGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCGCTCCCGACCGACCA 420
DB 361 TCGTGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCGCTCCCGACCGACCA 420

QY 421 ATACGACGCGCAGCTGATTCACAGCTGTTACCATCTCGCTTCGCGCGCATGACCGTG 480
DB 421 ATACGACGCGCAGCTGATTCACAGCTGTTACCATCTCGCTTCGCGCGCATGACCGTG 480

QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAAGGGTCAACGCTATGGCTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAAGGGTCAACGCTATGGCTGGGAT 540

QY 541 ATGATGATGAATGCTGGCTACCAAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATCTTC 600
DB 541 ATGATGATGAATGCTGGCTACCAAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATCTTC 600

QY 601 TAA 603
DB 601 CAA 603

RESULT 5
US-08-612-973-27
; Sequence 27, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..630
US-08-612-973-27

Query Match 98.6%; Score 597.4; DB 3; Length 636;
Best Local Similarity 99.8%; Pred. No. 1.7e-150;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGSCCTGGCGATGGCGTCCGGGTT 120
DB 61 CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGSCCTGGCGATGGCGTCCGGGTT 120

QY 121 CTGGAGGACGCGTGAACTATGCAACAGGGAATTTGCCGGTTGCTTTCTATCTTC 180
DB 121 CTGGAGGACGCGTGAACTATGCAACAGGGAATTTGCCGGTTGCTTTCTATCTTC 180

QY 181 CTCCTGGCTTTGCTGCTCTGCTGACCGTTCCAGTTCCCGTTATGAAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTGCTGCTCTGCTGACCGTTCCAGTTCCCGTTATGAAGTGGCAACGTG 240

QY 241 TCCGGGATGATACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300
DB 241 TCCGGGATGATACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300

QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAACACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAACACTCTTCCCGC 360

Db 301 GACATGATCATGACACCCCGGGTGGTGGCTTGGGAGAACAACTTTCCCGC 360
QY 361 TGTGGGTAGCGTCAACCCCGAGCTCGAGCTAGGAGCCAGCTCCCGACGACA 420
Db 361 TGTGGGTAGCGTCAACCCCGAGCTCGAGCTAGGAGCCAGCTCCCGACGACA 420
QY 421 ATAGGAGCGGCGTCACTGATTCGAGCTGTTACCATCTCGGCTCGCGGATGAGCGGTG 480
Db 421 ATAGGAGCGGCGTCACTGATTCGAGCTGTTACCATCTCGGCTCGCGGATGAGCGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAACCGGTCACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAACCGGTCACCGTATGGCTTGGGAT 540
QY 541 ATGATGATGAACCTGTCGCTCAACCGGCGCTGCTGATTCGAGCTGCTCCGATCGT 599
Db 541 ATGATGATGAACCTGTCGCTCAACCGGCGCTGCTGATTCGAGCTGCTCCGATCGT 599

RESULT 6

US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630

US-08-927-597-27

Query Match 98.6%; Score 597.4; DB 3; Length 636;
Best Local Similarity 99.8%; Pred. No. 1.7e-150;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAAGTCAATCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAAGTCAATCCCTTACATCGGGCTTCGCGACCTCGTGGGGTACATT 60
QY 61 CGCTCGTCCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTGCGCGATGCGCGTTC 120
Db 61 CGCTCGTCCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTGCGCGATGCGCGTTC 120
QY 121 CTGGAGGAGCGGCTGAACTATGCAACAGGGAATTTCCCGGTTGCTTCTTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACTATGCAACAGGGAATTTCCCGGTTGCTTCTTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGACCGGTTCCAGCTTCCGCTTATCAAGTGCGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGACCGGTTCCAGCTTATCAAGTGCGCAACGTG 240
QY 241 TCCGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300
Db 241 TCCGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGC 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTGCTGCGTTCGGGAGAACAACTTCTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGCTGCTGCGTTCGGGAGAACAACTTCTCCCGC 360
QY 361 TGTGGGTAGCGCTCAACCCCGAGCTCGAGCTAGGAGCCAGGCTCCCGACGACA 420
Db 361 TGTGGGTAGCGCTCAACCCCGAGCTCGAGCTAGGAGCCAGGCTCCCGACGACA 420
QY 421 ATACGACGCCACGCTCGATTCCAGCTGTTTACCATTCTCGCTCGCGGATGAGACGGTG 480
Db 421 ATACGACGCCACGCTCGATTCCAGCTGTTTACCATTCTCGCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGTCACCGTATGGCTTGGAT 540
QY 541 ATGATGATGAACCTGTCGCTCAACCGGCGCTGCTGATTCGAGCTGCTCCGATCCT 599
Db 541 ATGATGATGAACCTGTCGCTCAACCGGCGCTGCTGATTCGAGCTGCTCCGATCCT 599

RESULT 7

US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996

Mon Dec 22 13:28:42 2003

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555
; US-08-612-973-23

Query Match          91.7%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.8e-139;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEBERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555
; US-08-927-597-23

Query Match          91.7%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.8e-139;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGTTGGGTAAGGTCATCGATACCCCTTACATGGGGCTTCCCGACCTCGTGGGGTACATT 60
DB      1 ATGTTGGGTAAGGTCATCGATACCCCTTACATGGGGCTTCCCGACCTCGTGGGGTACATT 60

QY      61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB      61 CCGCTCGTCGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY      121 CTGAGGACGGCGTGAACTATGCAACAGGGAATTGGCCGGTTGCTCTTCTATCTTC 180
DB      121 CTGAGGACGGCGTGAACTATGCAACAGGGAATTGGCCGGTTGCTCTTCTATCTTC 180

QY      181 CTCTTGGCTTTGCTCTCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTG 240
DB      181 CTCTTGGCTTTGCTCTCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTG 240

QY      241 TCCGGGATGTACCATGTCACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB      241 TCCGGGATGTACCATGTCACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

QY      301 GACATGATCATGCACACCCCGGGTGGTGGCCCTCGTTCGGGAGAACAACTCTTCCCGC 360
DB      301 GACATGATCATGCACACCCCGGGTGGTGGCCCTCGTTCGGGAGAACAACTCTTCCCGC 360

QY      361 TCGTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGCAGCGTCCCAACACGACA 420
DB      361 TCGTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAACGCGCAGCGTCCCAACACGACA 420

QY      421 ATACGACGCGCAGTCCGATTCACAGTGTTCACATCTCGCTCGCCGCGCATGACGCGTG 480
DB      421 ATACGACGCGCAGTCCGATTCACAGTGTTCACATCTCGCTCGCCGCGCATGACGCGTG 480

QY      481 CAGGACTGCNAATTGCTCAATCTATCCCGGCGCATAAACGGGTCAACGATATGGCTTTGGGAT 540
DB      481 CAGGACTGCNAATTGCTCAATCTATCCCGGCGCATAAACGGGTCAACGATATGGCTTTGGGAT 540

QY      541 ATGATGATGAACCTGGT 556
DB      541 ATGATGATGAACCTGGT 556

RESULT 8
US-08-927-597-23
; Sequence 23, Application US/08927597

```

241	Db	TCGGGATGTACCATGTACGAAACGACTGCTTCCAACTCAAGCATTTGTATAGGCGAGCG	300
301	Qy	GACATGATCATGACACACCCCGGGTGCCTGCGTTGGGAGAAACAACCTCTTCCCGC	360
301	Db	GACATGATCATGACACACCCCGGGTGCCTGCGTTGGGAGAAACAACCTCTTCCCGC	360
361	Qy	TGCTGGGTAGCGCTCACCCCAACGCTCGCACTAGGAACGCCAGCGTCCCAACACGACA	420
361	Db	TGCTGGGTAGCGCTCACCCCAACGCTCGCACTAGGAACGCCAGCGTCCCAACACGACA	420
421	Qy	ATACAGCGCACGTGATTCACAGCTGTTACCACTATCTCGCTCGCGGCATGAGAGCGTG	480
421	Db	ATACAGCGCACGTGATTCACAGCTGTTACCACTATCTCGCTCGCGGCATGAGAGCGTG	480
481	Qy	CAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTACCGGTATGGGAT	540
481	Db	CAGGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTACCGGTATGGGAT	540
541	Qy	ATGATGATGAACGTGT	556
541	Db	ATGATGATGAACGTGT	556

RESULT 9
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

ADDRESS: NIXON & VANDERHYTE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/POCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5

LOCATION: 1..789
US-08-612-973-5

Query Match	85.2%	Score 516.2;	DB 3;	Length 795;
Best Local Similarity	88.9%	Pred. No. 8.5e-129;		
Matches 600;	Conservative	0;	Mismatches 3;	Indels 72

Qy	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAGGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTCGGCGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGGCATGGCGTCCGGGT	120
Db	61	CCGCTCGTCGGCGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGGCATGGCGTCCGGGT	120
Qy	121	CTGAGGACGGCGTGAACACTATGCAACAGGGAAATTTGGCCGGTGTCTTTCTCTATCTTC	180
Db	121	CTGAGGACGGCGTGAACACTATGCAACAGGGAAATTTGGCCGGTGTCTTTCTCTATCTTC	180
Qy	181	CTCTTGGCTTTGCTGCTCTGTGAACGGTTCACGTTCCGGTTATGAAGTCGGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGCTCTGTGAACGGTTCACGTTCCGGTTATGAAGTCGGCAACGTG	240
Qy	241	TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG	300
Db	241	TCCGGGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG	300
Qy	301	GACATGATCATGCACACCCCGGGTGGTGCCTCGTTCCGGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGCACACCCCGGGTGGTGCCTCGTTCCGGGAGAACAACTCTTCCCGC	360
Qy	361	TGCTGGGTAGCGGTCAACCCACGCTCGCAGCTAGGAACGCCAGCGTCCCAACACACGACA	420
Db	361	TGCTGGGTAGCGGTCAACCCACGCTCGCAGCTAGGAACGCCAGCGTCCCAACACACGACA	420
Qy	421	ATAAGACGCCACGTCGAT-----	438
Db	421	ATACGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG	480
Qy	439	-----TCCGAGCTGTTACCACTCGCCTCGCCGG	468
Db	481	GGGGAACCTTCGGGATCTGCTTCCTCGTCTCCGAGCTGTTACCATCTCGCCTCGCCGG	540
Qy	469	CATGAGACGGTCAGGACTGCAATTGCTCAATCTATCCGSGCCACATAACGGGTCAACCGT	528
Db	541	CATGAGACGGTCAGGACTGCAATTGCTCAATCTATCCGSGCCACATAACGGGTCAACCGT	600
Qy	529	ATGCGTTGGGATATGATGATGAACCTGGTCGCCTAACAGGGCCCTGGTGGTATCGCAGCTG	588
Db	601	ATGCGTTGGGATATGATGATGAACCTGGTCGCCTAACAGGGCCCTGGTGGTATCGCAGCTG	660
Qy	589	CTCCGGATCTCTAA	603
Db	661	CTCCGGATCCCAAA	675

RESULT 10
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, CEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA

STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM: 1.

QY 4 TTGGGTAAGTGCATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGTACATTCGG 63
DB 4 TTGGGTAAGTGCATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGTACATTCGG 63
QY 64 CTCGTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGCATGCGCTCGGGTTCTG 123
DB 64 CTCGTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGCATGCGCTCGGGTTCTG 123
QY 124 GAGGACGGCGTGAACATATCAACAGGAAATTTGCCGGTTGCTCTTCTCTATCTTCCTC 183
DB 124 GAGGACGGCGTGAACATATCAACAGGAAATTTGCCGGTTGCTCTTCTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGCTGCTGACCGCTTCAGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 243
DB 184 TTGGCTTTGCTGCTGCTGACCGCTTCAGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 243
QY 244 GGGATGTACATGTACAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCGGAC 303
DB 244 GGGATGTACATGTACAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCGGAC 303
QY 304 ATGATCATGCACACCCCGGGTGCCTGCTTCGGGAGAACTCTTCCCGCTGC 363
DB 304 ATGATCATGCACACCCCGGGTGCCTGCTTCGGGAGAACTCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTACCCCGAGCTCGAGCTAGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 423
DB 364 TGGGTAGCGCTACCCCGAGCTCGAGCTAGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 423
QY 424 CGACGCCAGCTCGAT----- 438
DB 424 CGACGCCAGCTCGAT----- 438
QY 439 -----TCCAGCTGTTCCACCATCTCGCTCGCGGCAAT 471
DB 484 GACCTCTCGGATCTGCTCTCTCCAGCTGTTCCACCATCTCGCTCGCGGCAAT 543
QY 472 GAGAGCTGCAGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGCTCACCGTATG 531
DB 544 GAGAGCTGCAGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGCTCACCGTATG 603
QY 532 GCTTGGGATATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
DB 604 GCTTGGGATATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY 592 CGGATCCTCTAA 603
DB 664 CGGATCCTCTAA 675

RESULT 12
US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..2076
US-08-927-597-47

Query Match 84.7%; Score 513.2; DB 3; Length 2082;
Best Local Similarity 88.8%; Pred. No. 7,1e-128;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;
QY 4 TTGGGTAAGTGCATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGTACATTCGG 63
DB 4 TTGGGTAAGTGCATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGTACATTCGG 63
QY 64 CTCGTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGCATGCGCTCGGGTTCTG 123
DB 64 CTCGTCGGCGCCCTAGGGGCGCTGCAGGGCCCTGCGCATGCGCTCGGGTTCTG 123
QY 124 GAGGACGGCGTGAACATATCAACAGGAAATTTGCCGGTTGCTCTTCTCTATCTTCCTC 183
DB 124 GAGGACGGCGTGAACATATCAACAGGAAATTTGCCGGTTGCTCTTCTCTATCTTCCTC 183
QY 184 TTGGCTTTGCTGCTGCTGACCGCTTCAGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 243
DB 184 TTGGCTTTGCTGCTGCTGACCGCTTCAGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 243
QY 244 GGGATGTACATGTACAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCGGAC 303
DB 244 GGGATGTACATGTACAGAACGACTGCTCCAACTCAAGCATTTGTGTATGAGGAGCGGAC 303
QY 304 ATGATCATGCACACCCCGGGTGCCTGCTTCGGGAGAACTCTTCCCGCTGC 363
DB 304 ATGATCATGCACACCCCGGGTGCCTGCTTCGGGAGAACTCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTACCCCGAGCTCGCTCGAGCTAGGAGCGGAGCGGAGCGGAGCGGAG 423
DB 364 TGGGTAGCGCTACCCCGAGCTCGCTCGAGCTAGGAGCGGAGCGGAGCGGAGCGGAG 423
QY 424 CGACGCCAGCTCGAT----- 438
DB 424 CGACGCCAGCTCGAT----- 438
QY 439 -----TCCAGCTGTTCCACCATCTCGCTCGCGGCAAT 471
DB 484 GACCTCTCGGATCTGCTCTCTCCAGCTGTTCCACCATCTCGCTCGCGGCAAT 543
QY 472 GAGAGCTGCAGACTGCAATTTGCTCAATCTATCCCGGCAATACCGGCTCACCGTATG 531

b	415	CTCGTCGGGCCCCCTAGGGGGCGCTGCACAGGGCCCTGGGCGATGGCGTTCGGGTTCTTG	474
y	124	GAGGAGCGGCTGAACATATGACACAGGGAATTTGCCGGTGTGCTCTTTCTCTATCTTCCTC	183
b	475	GAGGACGGGCGTGAACATATGACACAGGGAATTTGCCGGTGTGCTCTTTCTCTATCTTCCTC	534
y	184	TTGGCTTTTCTCTCTCTGTCTGACCGTTCCAGCTTCGGTTCGCTTATGAAGTGCGCAACGTGCC	243
b	535	TTGGCTTTTCTCTCTCTGTCTGACCGTTCCAGCTTCGGTTCGCTTATGAAGTGCGCAACGTGCC	594
y	244	GGGATGTAACATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGAC	303
b	595	GGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGGAC	654
y	304	ATGATCATGCACACCCCGGGTGGCTGCGCTTCGGTGAGAACAACTCTTCCCGCTGC	363
b	655	ATGATCATGCACACCCCGGGTGGCTGCGCTTCGGTGAGAACAACTCTTCCCGCTGC	714
y	364	TGGGTAGCGCTACCCCGACCGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACAATA	423
b	715	TGGGTAGCGCTACCCCGACCGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACAATA	774
y	424	CGACGCCACGTCGAT	438
b	775	CGACGCCACGTCGATTTGGGCGGCTGCTTCTGTTCGCTATGTACGTGGG	834
y	439	-----TCCAGCTGTTACCACTTCGCCTCGCGGCGAT	471
b	835	GACCTTCGGGATCTGCTTCCTCGTCTCCAGCTGTTACCACTTCGCCTCGCGGCGAT	894
y	472	GAGACGCTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATG	531
b	895	GAGACGCTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATG	954
y	532	GCTTGGGATATGATGAACTGCTCGCTTACAGCGCCCTGGTGGTATCGCAGCTGCTC	591
b	955	GCTTGGGATATGATGAACTGCTCGCTTACAGCGCCCTGGTGGTATCGCAGCTGCTC	1014
y	592	CGGATCCTCTAA 603	
b	1015	CGGATCCACAA 1026	
RESULT 14			
US-08-927-597-49			
; Sequence 49, Application US/08927597			
; Patent No. 6245503			
; GENERAL INFORMATION:			
; APPLICANT: MAERTENS, GEERT			
; APPLICANT: BOSMAN, PONS			
; APPLICANT: DE MARTYNOFF, GUY			
; APPLICANT: BUYSSE, MARIE-ANGE			
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE			
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE			
; NUMBER OF SEQUENCES: 111			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: NIXON & VANDERHVE P.C.			
; STREET: 1100 NORTH GLEBE ROAD			
; CITY: ARLINGTON			
; STATE: VIRGINIA			
; COUNTRY: U.S.A.			
; ZIP: 22201-4714			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/927,597			
; FILING DATE:			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			

APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-927-597-49

Query Match 84.7%; Score 513.2; DB 3; Length 2433;
Best Local Similarity 88.8%; Pred. No. 7.5e-128;
Matches 597; Conservative 0; Mismatches 3; Indels 72; Gaps 1;
QY 4 TTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 63
DB 355 TTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATTCCG 414
QY 64 CTGCTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTGCGCATGGCGTCCGGGTTCTG 123
DB 415 CTGCTCGGCGCCCCCTAGGGGCGCTGCCAGGCGCTGCGCATGGCGTCCGGGTTCTG 474
QY 124 GAGGACGGGTGAATCATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTCCTC 183
DB 475 GAGGACGGGTGAATCATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTCTGCTGACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTTCC 243
DB 535 TTGGCTTTGCTGCTCTGCTGACCGTTTCAGCTTCGCTTATGAAGTGGCAACGTTCC 594
QY 244 GGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 303
DB 595 GGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 654
QY 304 ATGATCATGCACACCCCGGTCGTCGCCCTCGGTTTCGGAGAACAACTCTTCCCGCTGC 363
DB 655 ATGATCATGCACACCCCGGTCGTCGCCCTCGGTTTCGGAGAACAACTCTTCCCGCTGC 714
QY 364 TGGGTAGCGTCAACCCGACGCTCGAGCTAGGAAAGCCAGCGTCCCGACCAACAATA 423
DB 715 TGGGTAGCGTCAACCCGACGCTCGAGCTAGGAAAGCCAGCGTCCCGACCAACAATA 774
QY 424 CGACGCCACGTCCAT----- 438
DB 775 CGACGCCACGTCCATTTGCTGTTGGGGGGGCTGCTTCTTCGCTATGTACGTGGG 834
QY 439 -----TCCAGCTGTTACCATCTCGCTTCGCTATGTACGTGGG 471
DB 835 GACCTCTGGGATCTGCTCTTCCTGCTCCAGCTGTTACCAATCTCGCTTCGCGGAT 894
QY 472 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACCGTATG 531
DB 895 GAGACGGTGCAGGACTGCAATTGCTCAATCTATCCGGCCACATAACGGGTCAACCGTATG 954
QY 532 GCTTGGGATATGATGATGAAGTGGTCCCTACAAAGCGCCCTGTTGTTATCGAGTCTCTC 591
DB 955 GCTTGGGATATGATGATGAAGTGGTCCCTACAAAGCGCCCTGTTGTTATCGAGTCTCTC 1014

QY 592 CGGATCTCTTAA 603
DB 1015 CGGATCCCAAA 1026
RESULT 15
US-08-612-973-7
Sequence 7, Application US/08612973
Patent No. 6150134
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 633 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..630
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..627
US-08-612-973-7
Query Match 75.3%; Score 456.4; DB 3; Length 633;
Best Local Similarity 86.8%; Pred. No. 7.4e-113;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT 60
QY 61 CGCTCGTCCGGCGCCCCCTAGGGGCGCTGCGAGGCGCTGCGCATGGCGTCCGGGTT 120
DB 61 CGCTCGTCCGGCGCCCCCTAGGGGCGCTGCGAGAGCCCTGCGCATGGCGTCCGGGTT 120
QY 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTTC 180

Mon Dec 22 13:28:42 2003

Qy	181	CTCTTGGCTTTGCTGTCTGTCTACACGGTTCAGCTTCCGCTTATGAAGTGCACGACGTG	240
Db	181	CTCTTGGCTTTACTGTCTGTCTTGACCACTTCAGCTTCGGCTTATGAGTGCACGACGTG	240
Qy	241	TCCGGGATGTAACCATGTGCAGAAAGACTGTCTCCAACTCAAGCATTTGTATGAGGCAGCG	300
Db	241	TCCGGGATGTACCATGTGCAGAAAGACTGTCTCCAACTCAAGCATTTGTATGAGGCAGCG	300
Qy	301	GACATGATCATGTGCACACCCCGGGTGCCTGCTTCGGGAGAAACAATCTTCCGCG	360
Db	301	GACATGATCATGTGCACACCCCGGGTGCCTGCTTCGGGAGAAACAATCTTCCGCG	360
Qy	361	TGCTGGGTAGCGCTCACCCGACGCTCGCAGTAGGAAGCCAGCGTCCCAACACGACA	420
Db	361	TGCTGGGTAGCGCTCACCCGACGCTCGCAGTAGGAAGCCAGCGTCCCACTACGACA	420
Qy	421	ATAACGACGCCAGCTCGAT-----	438
Db	421	ATAACGACGCCAGCTCGATTTGCTGTGGGGCGGCTGCTTCTGTTCGCTATGTACGTG	480
Qy	439	-----TCCGAGCTGTCACCATCTCGCTCGCGG	468
Db	481	GGGGATCTCTGGGATCTGTCCTTCGCTCTCCAGCTGTTACCAATCTCGCTCGCGG	540
Qy	469	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCAATAAACGGGTCAACGT	528
Db	541	CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCGCACATAACAGGTCAACGT	600
Qy	529	ATGGCTTGGGATATGATGAAGTACCTGGT	556
Db	601	ATGGCTTGGGATATGATGAAGTACCTGGT	628

Search completed: December 20, 2003, 07:03:12
Job time : 46.7045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2408.07 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303a-27
Perfect score: 636
Sequence: 1 ATGTTGGTAGGTCATCGA.....ATCACCACCACTATATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hcg_hum.*

31: em_hcg_inv.*

32: em_hcg_other.*

33: em_hcg_mus.*

34: em_hcg_pln.*

35: em_hcg_rod.*

36: em_hcg_mam.*

37: em_hcg_vrt.*

38: em_sy.*

39: em_higo_hum.*

40: em_higo_mus.*

41: em_higo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	636	100.0	636	6	A48689	A48689 Sequence 27
2	636	100.0	636	6	AR157339	AR157339 Sequence
3	636	100.0	636	6	AX452776	AX452776 Sequence
4	636	100.0	636	6	AX685028	AX685028 Sequence
5	597.4	93.9	606	6	A48687	A48687 Sequence 25
6	597.4	93.9	606	6	AR157338	AR157338 Sequence
7	597.4	93.9	606	6	AX452774	AX452774 Sequence
8	597.4	93.9	606	6	AX685026	AX685026 Sequence
9	597.4	93.9	723	6	A48683	A48683 Sequence 21
10	597.4	93.9	723	6	AR157336	AR157336 Sequence
11	597.4	93.9	723	6	AX452770	AX452770 Sequence
12	597.4	93.9	723	6	AX685022	AX685022 Sequence
13	556	87.4	561	6	A48685	A48685 Sequence 23
14	556	87.4	561	6	AR157337	AR157337 Sequence
15	556	87.4	561	6	AX452772	AX452772 Sequence
16	556	87.4	561	6	AX685024	AX685024 Sequence
17	515	81.0	795	6	A48667	A48667 Sequence 5
18	515	81.0	795	6	AR157325	AR157325 Sequence
19	515	81.0	795	6	AX452754	AX452754 Sequence
20	515	81.0	795	6	AX685006	AX685006 Sequence
21	512	80.5	2082	6	A48709	A48709 Sequence 47
22	512	80.5	2082	6	AR157350	AR157350 Sequence
23	512	80.5	2082	6	AX452796	AX452796 Sequence
24	512	80.5	2082	6	AX685048	AX685048 Sequence
25	512	80.5	2433	6	A48711	A48711 Sequence 49
26	512	80.5	2433	6	AR157351	AR157351 Sequence
27	512	80.5	2433	6	AX452798	AX452798 Sequence
28	512	80.5	2433	6	AX685050	AX685050 Sequence
29	459.2	72.2	3296	14	AB008441	AB008441 Hepatitis
30	459.2	72.2	3296	14	AB008442	AB008442 Hepatitis
31	459.2	72.2	3296	14	AB008443	AB008443 Hepatitis
32	459.2	72.2	9379	14	AF207766	AF207766 Hepatitis
33	456.4	71.8	633	6	A48669	A48669 Sequence 7
34	456.4	71.8	633	6	AR157326	AR157326 Sequence
35	456.4	71.8	633	6	AX452756	AX452756 Sequence
36	456.4	71.8	633	6	AX685008	AX685008 Sequence
37	456	71.7	9379	14	AF165052	AF165052 Hepatitis
38	456	71.7	9418	14	HCV132996	HCV132996 Hepatitis
39	454.4	71.4	3296	14	AB008445	AB008445 Hepatitis
40	454.4	71.4	3296	14	AB008446	AB008446 Hepatitis
41	454.4	71.4	9344	14	AB049096	AB049096 Hepatitis
42	454.4	71.4	9379	14	HCVPOLYP	HCVPOLYP Hepatitis
43	454.4	71.4	9386	14	AF165056	AF165056 Hepatitis
44	452.8	71.2	1615	14	HPCNS1SRJ	M74813 Hepatitis C
45	452.8	71.2	3296	14	AB008447	AB008447 Hepatitis

ALIGNMENTS

RESULT 1
A48689
LOCUS A48689 Sequence 27 from Patent WO9604385. 636 bp DNA linear PAT 07-MAR-1997
DEFINITION A48689
ACCESSION A48689
VERSION A48689.1 GI:2302402
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 636)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buyse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 27 15-FEB-1996;

LOCUS	AR157339	636 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 27 from patent US 6245503.				
ACCESSION	AR157339				
VERSION	AR157339.1	GI:16218273			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 636)				
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.				
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use				
JOURNAL	Patent: US 6245503-A 27 12-JUN-2001;				
FEATURES	Location/Qualifiers				
source	1..636				
BASE COUNT	119 a	203 c	174 g	140 t	
ORIGIN	1..636				
	Query Match 100.0%; Score 636; DB 6; Length 636;				
	Best Local Similarity 100.0%; Pred. No. 6.4e-130;				
	Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGTTGGGTAAGGTTCATCGATACCTTACATGGCGCTTCGCCGACCTCGTGGGGTACATT	60		
Db	1	ATGTTGGGTAAGGTTCATCGATACCTTACATGGCGCTTCGCCGACCTCGTGGGGTACATT	60		
Qy	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGTGCAGAGGCCCTGGCGCATGGCGTCCGGGT	120		
Db	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGTGCAGAGGCCCTGGCGCATGGCGTCCGGGT	120		
Qy	121	CTGAGGACGCGCGTGAATATGCAACAGGGAAATTGGCCGGTTGCTTTCTATCTTC	180		
Db	121	CTGAGGACGCGCGTGAATATGCAACAGGGAAATTGGCCGGTTGCTTTCTATCTTC	180		
Qy	181	CTCTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG	240		
Db	181	CTCTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG	240		
Qy	241	TCCGGGATGATACCATGTCCAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG	300		
Db	241	TCCGGGATGATACCATGTCCAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG	300		
Qy	301	GACATGATCATGCACACCCCCGGGTGGTCCCTCGCTTCGGGAGAACACTCTTCCCGC	360		
Db	301	GACATGATCATGCACACCCCCGGGTGGTCCCTCGCTTCGGGAGAACACTCTTCCCGC	360		
Qy	361	TGCTGGGTAGCGCTCACCCCGCTCCAGCTCGCAGCTAGGAAACGCGTCCCAACGACA	420		
Db	361	TGCTGGGTAGCGCTCACCCCGCTCCAGCTCGCAGCTAGGAAACGCGTCCCAACGACA	420		
Qy	421	ATACGACGCGCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	480		
Db	421	ATACGACGCGCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	480		
Qy	481	CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGGTACCGTATGGGAT	540		
Db	481	CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACCGGGTACCGTATGGGAT	540		
Qy	541	ATGATGATGAACTGGTTCGCTACACGGCCCTTGGTGGTATTCGACGCTGCTCCGGATCGTG	600		
Db	541	ATGATGATGAACTGGTTCGCTACACGGCCCTTGGTGGTATTCGACGCTGCTCCGGATCGTG	600		
Qy	601	ATCGAGGCGACACCATCACCATCACTAATAG	636		
Db	601	ATCGAGGCGACACCATCACCATCACTAATAG	636		
RESULT 3					
AX452776					
LOCUS					
DEFINITION	Sequence 27 from Patent EP1211315.				
ACCESSION	AX452776	636 bp	DNA	linear	PAT 06-JUL-2002

COMMENT	INNOGENETICS NV (BE)			
	Other publication CA 2172273 960215			
	Other publication AU 3382495 960304.			
	Location/Qualifiers			
FEATURES	source	1..636		
		/organism="unidentified"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:32644"		
CDS		1..633		
		/note="unnamed protein product"		
		/codon_start=1		
		/protein_id="CAA03137.1"		
		/db_xref="GI:2302403"		
		/translations="MIGKVIDTITCGPADIVGPIVGPAPLGAARALAHGVRLVDG VNTATGNLPCQSFISIFLLALLSCLTLPASAYEVNYSVMYVNDSCNSISIVPEADAM IMHTPGCPVPCVRNNSRQVATPTLAARNASVPTTIRRHVDLSQLTFSIPRRHETY QDCNCSIYPGHITGHRMWMNMNNSPTTALVVSQLLRIVIEGRHHHHH"		
	mat_peptide	1..630		
		/product="unnamed"		
BASE COUNT	119 a	203 c	174 g	140 t
ORIGIN				
	Query Match	100.0%;	Score 636;	DB 6; Length 636;
	Best Local Similarity	100.0%;	Pred. No. 6.4e-130;	
	Matches 636; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGTTGGGTAAGGTTCATCGATACCTTACATGGCGCTTCGCCGACCTCGTGGGGTACATT	60	
DB	1	ATGTTGGGTAAGGTTCATCGATACCTTACATGGCGCTTCGCCGACCTCGTGGGGTACATT	60	
QY	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGGT	120	
DB	61	CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGCGCTTGGCGCATGGCGTCCGGGT	120	
QY	121	CTGAGGACGCGGTGAACATATGCAACAGGGAAATTGGCCGGTTGCTTTCTATCTTC	180	
DB	121	CTGAGGACGCGGTGAACATATGCAACAGGGAAATTGGCCGGTTGCTTTCTATCTTC	180	
QY	181	CTCTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG	240	
DB	181	CTCTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACACGTG	240	
QY	241	TCCGGGATGATACCATGTCCAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG	300	
DB	241	TCCGGGATGATACCATGTCCAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGCAGCG	300	
QY	301	GACATGATCATGCACACCCCGGGTGGTCCCTCGCTTCGGGAGAACACTCTTCCCGC	360	
DB	301	GACATGATCATGCACACCCCGGGTGGTCCCTCGCTTCGGGAGAACACTCTTCCCGC	360	
QY	361	TGCTGGGTAGCGCTCACCCCGCTCCAGCTCGCAGCTAGGAAACGCCAGCGTCCCAACGACA	420	
DB	361	TGCTGGGTAGCGCTCACCCCGCTCCAGCTCGCAGCTAGGAAACGCCAGCGTCCCAACGACA	420	
QY	421	ATACGACGCGACGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	480	
DB	421	ATACGACGCGACGTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	480	
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGGAT	540	
DB	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACCGTATGGGAT	540	
QY	541	ATGATGATGAACCTGGTTCGCTTACAAAGCGCCCTGGTGGTATTCGACGCTGCTCCGGAT	600	
DB	541	ATGATGATGAACCTGGTTCGCTTACAAAGCGCCCTGGTGGTATTCGACGCTGCTCCGGAT	600	
QY	601	ATCGAGGCGACACCATCACCATCACTAATAG	636	
DB	601	ATCGAGGCGACACCATCACCATCACTAATAG	636	
RESULT 2				

QY	421	ATACAGCCGACGTCGATCCAGCTGTTCCAGCTGTTCCAGCTGCTCCGCTCCGCGCATGAGACGGTG	480
Db	421	ATACAGCCGACGTCGATCCAGCTGTTCCAGCTGTTCCAGCTGCTCCGCTCCGCGCATGAGACGGTG	480
QY	481	CAGGACTCAATGCTCAATCTATCCCGCCACATAACGGGTACCCGTTGCGGTATGGCTTGGGAT	540
Db	481	CAGGACTCAATGCTCAATCTATCCCGCCACATAACGGGTACCCGTTGCGGTATGGCTTGGGAT	540
QY	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGTTGGTATGCCAGCTGCTCCGATCGTG	600
Db	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGTTGGTATGCCAGCTGCTCCGATCGTG	600
QY	601	ATCGAGGCGAGACACCATCACCACCATCACTAATAG	636
Db	601	ATCGAGGCGAGACACCATCACCACCATCACTAATAG	636
RESULT 5			
LOCUS	A48687	606 bp	DNA linear PAT 07-MAR-1997
DEFINITION	Sequence 25 from Patent WO9604385.		
ACCESSION	A48687		
VERSION	A48687.1	GI:2302400	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 606)		
AUTHORS	Maertens, G., Bosman, F., De, M.G. and Buyse, M.		
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE		
JOURNAL	PATENT: WO 9604385-A 25 15-FEB-1996;		
COMMENT	INNOGENETICS NV (BE)		
OTHER PUBLICATION	CA 2172273 960215		
OTHER PUBLICATION	US 3382495 960304.		
FEATURES	Location/Qualifiers		
source	1..606		
/organism="unidentified"			
/mol_type="genomic DNA"			
/db_xref="taxon:32644"			
1..603			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAA03136.1"			
/db_xref="GI:2302401"			
/translation="MLGKVIDTLTCGFADLVGYPLVGAPLGGARALAHGVRLVLEDDVNTATNLPCSPSIFLLALLSLCTVPASAVEVNVSMVHVTNDCNSISVVEAADMIMHTPGVCVPCVRENNSSRCWALTPILARNASVPTTIRRHVDSQLFTISPRRHETVQDCNCSTYPGHITGHRMADMMNMNSPTTALVVSQLLRIL"			
1..600			
/product="unnamed"			
mat_peptide	109 a	193 c	167 g 137 t
BASE COUNT			
ORIGIN			
Query Match	93.9%;	Score 597.4;	DB 6; Length 606;
Best Local Similarity	99.8%;	Pred. No. 2e-121;	
Matches 598;	Conservative	0; Mismatches	1; Indels 0; Gaps 0;
QY	1	ATGTTGGGTAGGTGATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT	60
Db	1	ATGTTGGGTAGGTGATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT	60
QY	61	CGGCTCGTCCGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGGCGTCCGGGTT	120
Db	61	CGGCTCGTCCGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGGCGTCCGGGTT	120
QY	121	CTGAGGACGGGTGAATCTGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC	180
Db	121	CTGAGGACGGGTGAATCTGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGGCTTATGAGTGGCGAACGTTG	240
Db	181	CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGGCTTATGAGTGGCGAACGTTG	240
QY	241	TCGGGATGTACATGTGCAAGACGACTGCTCCACTCAAGCATTTGTGTATGAGGACGG	300
Db	241	TCGGGATGTACATGTGCAAGACGACTGCTCCACTCAAGCATTTGTGTATGAGGACGG	300
QY	301	GACATGATCATGACACACCCCGGCTGCTGCTTCGGTTCGGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGACACACCCCGGCTGCTGCTTCGGTTCGGGAGAACAACTCTTCCCGC	360
QY	361	TGCTGGGTAGCGCTCACACCCCGGCTGCGAGCTAGGAACCGCAGCGTCCCAACGACGACA	420
Db	361	TGCTGGGTAGCGCTCACACCCCGGCTGCGAGCTAGGAACCGCAGCGTCCCAACGACGACA	420
QY	421	ATACGACGCGACGCTGATTTCCAGCTGTTTCAATCTCCGCTCGCGGATGAGACGGTG	480
Db	421	ATACGACGCGACGCTGATTTCCAGCTGTTTCAATCTCCGCTCGCGGATGAGACGGTG	480
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACGGGTACCCGTTGCGGTATGGGAT	540
Db	481	CAGGACTGCAATTTGCTCAATCTATCCCGCCACATAACGGGTACCCGTTGCGGTATGGGAT	540
QY	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGTTGGTATGCCAGCTGCTCCGATCGTG	599
Db	541	ATGATGATGAATGCTGCGCTACAAACGGCCCTGTTGGTATGCCAGCTGCTCCGATCGTG	599
RESULT 6			
LOCUS	AR157338	606 bp	DNA linear PAT 17-OCT-2001
DEFINITION	Sequence 25 from patent US 6245503.		
ACCESSION	AR157338		
VERSION	AR157338.1	GI:16218271	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 606)		
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.		
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use		
JOURNAL	Patent: US 6245503-A 25 12-JUN-2001;		
FEATURES	Location/Qualifiers		
source	1..606		
/organism="unknown"			
BASE COUNT	109 a	193 c	167 g 137 t
ORIGIN			
Query Match	93.9%;	Score 597.4;	DB 6; Length 606;
Best Local Similarity	99.8%;	Pred. No. 2e-121;	
Matches 598;	Conservative	0; Mismatches	1; Indels 0; Gaps 0;
QY	1	ATGTTGGGTAGGTGATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT	60
Db	1	ATGTTGGGTAGGTGATCATGATACCTTACATGCGGCTTCGCCGACCTCGTGGGTACATT	60
QY	61	CGGCTCGTCCGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGGCGTCCGGGTT	120
Db	61	CGGCTCGTCCGCGCCCTAGGGGCGCTGCGAGGGCCCTGCGCATGGCGTCCGGGTT	120
QY	121	CTGAGGACGGGTGAATCTGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC	180
Db	121	CTGAGGACGGGTGAATCTGCAACAGGGAATTTGCCGGTTGCTCTTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGGCTTATGAGTGGCGAACGTTG	240
Db	181	CTCTTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGGCTTATGAGTGGCGAACGTTG	240
QY	241	TCGGGATGTACATGTGCAAGACGACTGCTCCACTCAAGCATTTGTGTATGAGGACGG	300
Db	241	TCGGGATGTACATGTGCAAGACGACTGCTCCACTCAAGCATTTGTGTATGAGGACGG	300
QY	301	GACATGATCATGACACACCCCGGCTGCTGCTTCGGTTCGGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGACACACCCCGGCTGCTGCTTCGGTTCGGGAGAACAACTCTTCCCGC	360

Db	301	GACATGATCATGCACACCCCGGGTGGTGCCCTTGGGAGAACACTCTTCCCGC	360
Qy	361	TGCTGGGTAGCGCTCACCCACACGCTCGCAGCTAGGAACGCCAGCGTCCCACCACGACA	420
Db	361	TGCTGGGTAGCGCTCACCCACACGCTCGCAGCTAGGAACGCCAGCGTCCCACCACGACA	420
Qy	421	ATACGACGCCACGTCGATTCGCCAGCTGTTTCCACATCTCGGCTCGCCGGCATGAGACGGT	480
Db	421	ATACGACGCCACGTCGATTCGCCAGCTGTTTCCACATCTCGGCTCGCCGGCATGAGACGGT	480
Qy	481	CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT	540
Db	481	CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT	540
Qy	541	ATGATGATGAACGTGTCGCCCTACAAACGGCCCTCGTGGTATCGCAGCTGCCTCCGATCGT	599
Db	541	ATGATGATGAACGTGTCGCCCTACAAACGGCCCTCGTGGTATCGCAGCTGCCTCCGATCGT	599
RESULT 7			
AX452774			
LOCUS	AX452774	606 bp	linear
DEFINITION	Sequence 25 from Patent EP1211315.		PAT 06-JUL-2002
ACCESSION	AX452774		
VERSION	AX452774.1	GI:21712459	
KEYWORDS			
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE			
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buysse,M.A.		
TITLE	Recombinant vectors for producing hcv envelope proteins		
JOURNAL	Patent: Ep 1211315-A 25 05-JUN-2002;		
	Innogenetics N.V. (BE)		
FEATURES			
source	Location/Qualifiers		
	1..606		
	/organism="Hepatitis C virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:11103"		
CDS	1..603		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD38074.1"		
	/db_xref="GI:21712460"		
	/translation="MLGKVIDLTTCGPADLVGVIPLVGAPLPGNARALAHGVRLVLEGG		
	VNYATGNLPCFSIFILLALLSLCTVPASAYEVNRVSGMTHVNDCSNSIVYEADAM		
	IMHTPGCVPCVRENSRSRWALPTLAAENASVPTTIRRHVDSQLFTISPRRHET		
	QDCNCSYPIGHITGHRMAWMNMNWSPTTALVVSQLLRL"		
mat_peptide	1..600		
	/product="unnamed"		
BASE COUNT	109 a 193 c 167 g	137 t	
ORIGIN			
Query Match	93.9%	Score 597.4;	DB 6; Length 606;
Best Local Similarity	99.8%;	Pred. No. 2e-121;	
Matches	598;	Conservative 0;	Mismatches 1; Indels 0; Gaps 0;
Qy	1	ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAAGTCATCGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Db	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
Qy	121	CTGGAGGACGGCGTGAACATAATGCAACAGGAAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAACATAATGCAACAGGAAATTTGCCCGGTTGCTCTTTCTCTATCTTC	180
Qy	181	CTCTTGCTTGTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACGCGT	240

Db	181	CTCTTGCTTTGCTGTCCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTG	240
Qy	241	TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGCAGCG	300
Db	241	TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTTGTATGAGGCAGCG	300
Qy	301	GACATGATCATGACACACCCCGGGTGGTGCCTCGGAGAACAACTCTTCCCGC	360
Db	301	GACATGATCATGACACACCCCGGGTGGTGCCTCGGAGAACAACTCTTCCCGC	360
Qy	361	TGCTGGGTAGGCTCACCCCGAGCTCGAGCTAGGACGCCAGCGTCCGCCACACGACA	420
Db	361	TGCTGGGTAGGCTCACCCCGAGCTCGAGCTAGGACGCCAGCGTCCGCCACACGACA	420
Qy	421	ATACGAGCCACGTCGATTCGCCAGCTGTTCCACCATCTCGCCTCGCCGGCATGAGACGGTG	480
Db	421	ATACGAGCCACGTCGATTCGCCAGCTGTTCCACCATCTCGCCTCGCCGGCATGAGACGGTG	480
Qy	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540
Db	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT	540
Qy	541	ATGATGATGAACCTGCTGCCTACACGGCCCTGGTGTATCGAGCTGCTCCCGATCGT	599
Db	541	ATGATGATGAACCTGCTGCCTACACGGCCCTGGTGTATCGAGCTGCTCCCGATCGT	599
RESULT 8			
AX685026			
LOCUS	AX685026	606 bp	DNA
DEFINITION	Sequence 25 from Patent WO0205548.		linear
ACCESSION	AX685026		PAT 29-MAR-2003
VERSION	AX685026.1		
KEYWORDS	GI:29371431		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1		
AUTHORS	Maertens, G., Boeman, F. and Buyse, M.A.		
TITLE	Purified Hepatitis C Virus envelope proteins for diagnostic and therapeutic use		
JOURNAL	Patent: WO 0205548-A 25 18-JUL-2002;		
FEATURES	INNOGENETICS N.V. (BE)		
source	Location/Qualifiers		
	1..606		
	/organism="Hepatitis C virus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:11103"		
CDS	1..603		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD86527.1"		
	/db_xref="GI:29371432"		
	/translation="MLGKVIDTLTCGFADIVGLVPGAPLGGARALAHGVRVLGDD		
	VNYATGNLPCQSFIFLLALLSLCTVPASAYEVNVMYVTDNCSNLSIVVEADAM		
	IMHTPGCPVCRNNSRCWVALPTLAARNASVPTTIRRHVDSQLFTISPRRHETV		
	QDCNCSIRYPGHITGHRMADMMNNSPTALVVSQLLRLI"		
mat_peptide	1..600		
	/product="unnamed"		
BASE COUNT	109 a 193 c 167 g 137 t		
ORIGIN			
Query Match	93.9%	Score 597.4;	DB 6; Length 606;
Best local Similarity	99.8%	Pred. No. 2e-121;	
Matches	598; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	1	ATGTTGGTAAAGTTCATCGATACCCCTACATCGGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGTAAAGTTCATCGATACCCCTACATCGGCTTCGCCGACCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGCTCCGGGTT	120

Db	61	CGGCTCGTCGGCGCCCGCCCTAGGGGGCGCTGCCAGGCGCCCTGGCGCATGGCGTTCGGGTT	120
Qy	121	CTGGAGACGGCGTGAATCTATGCAACAGGAATTTGCCGGTTCGCTCTTTCTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAATCTATGCAACAGGAATTTGCCGGTTCGCTCTTTCTCTATCTTC	180
Qy	181	CTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCGCGTTATGAAAGTGGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCGCGTTATGAAAGTGGCAACGTG	240
Qy	241	TCCGGGATGTACATGTCAAGTCAAGACACTGCTCCAACTCAAGCATTTGTATGATGAGGCAGCG	300
Db	241	TCCGGGATGTACATGTCAAGAACACTGCTCCAACTCAAGCATTTGTATGATGAGGCAGCG	300
Qy	301	GACATGATCATGACACCCCGGGTGCCTGCGTTCCGGAGAACACTCTTCCCGC	360
Db	301	GACATGATCATGACACCCCGGGTGCCTGCGTTCCGGAGAACACTCTTCCCGC	360
Qy	361	TGCTGGGTAGCGTCAACCCCGCCAGCTCGAGCTAGGAAAGCGCAGCGTCCCCCACCACGACA	420
Db	361	TGCTGGGTAGCGTCAACCCCGCCAGCTCGGAGTAGGAAAGCGCAGCGTCCCCCACCACGACA	420
Qy	421	ATACGAGCCACGTCGATTCCCAAGCTGTTCAACATCTCGCTCGCCGGCATGACAGCGTG	480
Db	421	ATA CGA CGGCA CGTTCGATTCCCAAGCTGTTCAACATCTCGCTCGCCGGCATGACAGCGTG	480
Qy	481	CAGGACTGCAATTGCTCAATCTATCCGGGCGCATAAAGGGTCAACGTTATGGCTTGGGAT	540
Db	481	CAGGACTGCAATTGCTCAANTCTATCCGGGCGCATAAAGGGHTACCGTTATGGCTTGGGAT	540
Qy	541	ATGATGATGAACCTGGTGCCTCTACAAAGCGCCCTGGTGGTATCGCAGCTGCTCGGATCGT	599
Db	541	ATGATGATGAACCTGGTGCCTCTACAAAGCGCCCTGGTGGTATCGCAGCTGCTCGGATCGT	599

RESULT	9
A48683	
LOCUS	723 bp DNA linear PAT 07-MAR-1997
DEFINITION	Sequence 21 from Patent WO9604385.
ACCESSION	A48683
VERSION	A48683.1 GI:2302396
KEYWORDS	.
SOURCE	unidentified
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 723)
AUTHORS	Maertens,G., Bosman,F., De,M.G. and Buysse,M.
TITLE	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL	Patent: WO 9604385-A 21 15-FEB-1996;
COMMENT	INNOGENETICS NV (BE) Other publication CA 2172273 960215 Other publication AU 3382495 960304.
FEATURES	Location/Qualifiers
source	1..723 /organism="unidentified" /mol type="genomic DNA" /db xref="taxon:32644"
CDS	1..720 /note="unnamed protein product" /codon_start=1 /protein_id="CA03134.1" /db_xref="GI:2302397" /translation="MLGKVIDTLTCTGFADILVGYIPLVCGAPLGGNAALAHGVRLVDG VNYATGNLRGCSFSIFLLALLSLCTVPASAYEVNRVMGHVTNDSCNSSIYVEAADM IMHTPGCVPCVRNNRSRCWVALTPLAARNASVPTTIIRRHVDQSQLTFTISPRRHTY QDCNCSTYPGHITGHRMADMMNMNSPTTALVVSQLEIPQAVVDMVGAWHGVLGLA AYYSMVGNWAKVLIVMLLPAP"
mat_peptide	1..717 /product="unnamed"
BASE COUNT	. 126 a 220 c 208 g 169 t
ORIGIN	

Query Match	93.9%	Score 597	DB 6	Length 723
Best Local Similarity	100.0%	Pred. No. 2.5e-121	Indels	Gaps
Matches 597	Conservative	0	Mismatches 0	
1	ATCTGGGTAAGTCAATCGATACCCCTTACATGCGGTTCGCCGACTCGTGGGTACATT	60		
1	ATGTTGGGTAAGTCAATCGATACCCCTTACATGCGGTTCGCCGACTCGTGGGTACATT	60		
61	CCGCTCGTTCGGCGCCCTTAGGGGGCGTGCACAGGGCCCTGCGCGATGGCGTCCGGGTT	120		
61	CCGCTCGTTCGGCGCCCTTAGGGGGCGTGCACAGGGCCCTGCGCGATGGCGTCCGGGTT	120		
121	CTGGAGACGCGGTGAACCTATGCAACAGGGAATTTGCCGGTTGCTTCTCTATCTTC	180		
121	CTGGAGACGCGGTGAACCTATGCAACAGGGAATTTGCCGGTTGCTTCTCTATCTTC	180		
181	CTCTTGGCTTTGCTCCTCTCTGACCGTTTCCAGCTTTCGCGTTTATGAAGTGCACAAGTG	240		
181	CTCTTGGCTTTGCTCCTCTCTGACCGTTTCCAGCTTTCGCGTTTATGAAGTGCACAAGTG	240		
241	TCCGGGATGTACCATGTCAAGAACGATGTCTCCAACTCAAGCATTTGTATGAGGACGCG	300		
241	TCCGGGATGTACCATGTCAAGAACGATGTCTCCAACTCAAGCATTTGTATGAGGACGCG	300		
301	GACATGATCATGACACACCCCGGGTGGGTGCGCTTTCGGGAGAACAACTCTTCCGCG	360		
301	GACATGATCATGACACACCCCGGGTGGGTGCGCTTTCGGGAGAACAACTCTTCCGCG	360		
361	TGCTGGGTAGCGCTCAACCCCAACGCTGCGAGCTAGGAACGCCAGCGTCCCAACACGACA	420		
361	TGCTGGGTAGCGCTCAACCCCAACGCTGCGAGCTAGGAACGCCAGCGTCCCAACACGACA	420		
421	ATACACACGCGACGTCGATTTCCAGCTGTTCACCATCTCGCTTCGCGGCATGAGACGGTG	480		
421	ATACACGCGCACGTCGATTTCCAGCTGTTCACCATCTCGCTTCGCGGCATGAGACGGTG	480		
481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAACGGGTTCACCGTATGGCTTGGGAT	540		
481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAACGGGTTCACCGTATGGCTTGGGAT	540		
541	ATGATGATGAACGGTTCGCGCTTCAACCGGCCCTTGGTGGTATCGACGCTGCTCCGGATC	597		
541	ATGATGATGAACGGTTCGCGCTTCAACCGGCCCTTGGTGGTATCGACGCTGCTCCGGATC	597		

RESULT	10
AR157336	
LOCUS	linear DNA PAT 17-OCT-2001
DEFINITION	Sequence 21 from patent US 6245503.
ACCESSION	AR157336
VERSION	AR157336.1 GI:16218269
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 723)
TITLE	Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A. Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL	Patent: US 6245503-A 21 12-JUN-2001;
FEATURES	Location/Qualifiers
source	1..723
BASE COUNT	136 a 220 c 208 g 169 t
ORIGIN	/organism="unknown"
Query Match	93.9%; Score 597; DB 6; Length 723;
Best Local Similarity	100.0%; Pred. No. 2.5e+121;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 ATGTTGGGTAAAGGCATCGATTACCCTTACATCGCGCTTCGCCGACCTCGTGCGGTACATT 60
Dd	1 ATGTTGGGTAAAGGCATCGATTACCCTTACATCGCGCTTCGCCGACCTCGTGCGGTACATT 60

```
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTT 120
QY 121 CTGAGGAGCGGCGTGAATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
DB 121 CTGAGGAGCGGCGTGAATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGCTGCTGTCAGCCGTTCCAGCTTCCGCTTATGAAGTGGCGCAACGTTG 240
DB 181 CTCCTGGCTTTGCTGCTGCTGTCAGCCGTTCCAGCTTCCGCTTATGAAGTGGCGCAACGTTG 240
QY 241 TCCGGGATGACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGAGCG 300
DB 241 TCCGGGATGACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCCAGCTGATCCAGCTGTTCCAGCTTCCGCTGCGCGGATGAGACGGTG 480
DB 421 ATACGACGCCAGCTGATCCAGCTGTTCCAGCTTCCGCTGCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGGTCACGGTATGCTTGGGAT 540
DB 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGGTCACGGTATGCTTGGGAT 540
QY 541 ATGATGATGAATGGTTCGCTACACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATC 597
DB 541 ATGATGATGAATGGTTCGCTACACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATC 597

RESULT 11
AX452770
LOCUS AX452770 723 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 21 from Patent EP1211315.
ACCESSION AX452770
VERSION AX452770.1 GI:21712455
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins
JOURNAL Patent: EP 1211315-A 21 05-JUN-2002;
Innogenetics N.V. (BE)
FEATURES
source
1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38072.1"
/db_xref="GI:21712456"
/translation="MLGKVIDLTGCFADLVGYIFLVGAPLGGAARALAHGVRLVDG
VNYATGNLPGCSFISPLALLSLVTPASAYEVNVMGYHVNDCNSNIVYEADM
INHTPCVPCVENNSRCWALPTPLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNGSIVPGHITGHRMADMMNNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGL
AYISMVGNWAKVLIVMLLPAP"
mat_peptide 1..717
/product="unnamed"
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN
```

```
Query Match 93.9%; Score 597; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.5e-121;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCCGCCGACTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCCGCCGACTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTCGCGCATGGCGTCCGGGTT 120
QY 121 CTGAGGAGCGGCGTGAATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
DB 121 CTGAGGAGCGGCGTGAATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGCTGCTGTCAGCCGTTCCAGCTTCCGCTTATGAAGTGGCGCAACGTTG 240
DB 181 CTCCTGGCTTTGCTGCTGCTGTCAGCCGTTCCAGCTTCCGCTTATGAAGTGGCGCAACGTTG 240
QY 241 TCCGGGATGACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGAGCG 300
DB 241 TCCGGGATGACCATGTACAGAACGACTGCTCCAACCTCAAGCATTTGTGTATGAGGAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCGGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGGCTGCGAGTAGGAACGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCCAGCTGATCCAGCTGTTCCAGCTTCCGCTGCGCGGATGAGACGGTG 480
DB 421 ATACGACGCCAGCTGATCCAGCTGTTCCAGCTTCCGCTGCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGGTCACGGTATGCTTGGGAT 540
DB 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATACAGGGTCACGGTATGCTTGGGAT 540
QY 541 ATGATGATGAATGGTTCGCTACACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATC 597
DB 541 ATGATGATGAATGGTTCGCTACACAGCGCCCTGGTGTATCGCAGCTGCTCCGGATC 597

RESULT 12
AX685022
LOCUS AX685022 723 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 21 from Patent WO0205548.
ACCESSION AX685022
VERSION AX685022.1 GI:29371427
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Maertens,G., Bosman,F. and Buyse,M.A.
TITLE Purified Hepatitis C Virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 0205548-A 21 18-JUL-2002;
INNOGENETICS N.V. (BE)
FEATURES
source
1..723
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
1..720
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86525.1"
/db_xref="GI:29371428"
CDS
```

/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLBDG
VNYATGNLPGCSFIFLLALLSCLTLPASAVEYRVNMGVHYVNDGNSIIVYEAADM
IMHTPGCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGHITGHRMAWMDMMNNP
AYSNVGNWAKVLIVMLLFP"

mat_peptide 1..717
BASE COUNT 126 a 220 c 208 g 169 t
ORIGIN /product="unnamed"

Query Match 93.9%; Score 597; DB 6; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.5e-121;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCACCTCGTGGGGTACATT 60

QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGGCTTATGAAGTGCACAACGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGGCTTATGAAGTGCACAACGTG 240

QY 241 TCCGGGATGATACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGATACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGACGG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCGACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCGACGACA 420

QY 421 ATACAGCCACGCTCGATTCACGCTTCCAGCTTCGCGTCCGGCATGAGACGGTG 480
DB 421 ATACAGCCACGCTCGATTCACGCTTCCAGCTTCGCGTCCGGCATGAGACGGTG 480

QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

QY 541 ATGATGATGAATGCTCGCTTACAAAGCCGCTGCTGATTCGAGCTGCTCCGGATC 597
DB 541 ATGATGATGAATGCTCGCTTACAAAGCCGCTGCTGATTCGAGCTGCTCCGGATC 597

RESULT 13
A48685
LOCUS A48685 561 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 23 from Patent WO9604385.
ACCESSION A48685
VERSION A48685.1 GI:2302398
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified

REFERENCE 1 (bases 1 to 561)
AUTHORS Maertens, G., Bosman, F., De M.G. and Buysse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 23 15-FEB-1996;
INNOGENETICS NV (BE)
COMMENT Other publication CA 2172273 960215
Other publication AU 3382495 960304.

Location/Qualifiers
1..561
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1..558
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03135.1"
/db_xref="GI:2302398"
/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRVLBDG
VNYATGNLPGCSFIFLLALLSCLTLPASAVEYRVNMGVHYVNDGNSIIVYEAADM
IMHTPGCVPCVRENNSSRCWVALTPTLAARNASVPTTIRRHVDSQLFTISPRRHETV
QDCNCSIYPGHITGHRMAWMDMMNNP"
mat_peptide 1..555
BASE COUNT 103 a 176 c 155 g 127 t
ORIGIN /product="unnamed"

Query Match 87.4%; Score 556; DB 6; Length 561;
Best Local Similarity 100.0%; Pred. No. 2.7e-112;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCACCTCGTGGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCACCTCGTGGGGTACATT 60

QY 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAATATGCAACAGGGAATTTGCCGGTTCCTTCTCTATCTTC 180

QY 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGGCTTATGAAGTGCACAACGTG 240
DB 181 CTCTTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGGCTTATGAAGTGCACAACGTG 240

QY 241 TCCGGGATGATACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGATACCATGTCAAGAACGACTGTCTCAACTCAAGCATTTGTATGAGGACGG 300

QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTTCCGGAGAACAACTCTTCCCGC 360

QY 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCGACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCGACGACA 420

QY 421 ATACAGCCACGCTCGATTCACGCTTCCAGCTTCGCGTCCGGCATGAGACGGTG 480
DB 421 ATACAGCCACGCTCGATTCACGCTTCCAGCTTCGCGTCCGGCATGAGACGGTG 480

QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540
DB 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGGTATGGCTTGGGAT 540

QY 541 ATGATGATGAATGCTG 556
DB 541 ATGATGATGAATGCTG 556

RESULT 14
AR157337
LOCUS AR157337 561 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 23 from patent US 6245503.
ACCESSION AR157337
VERSION AR157337.1 GI:16218270
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.		Location/Qualifiers	
REFERENCE	1 (bases 1 to 561)	1..561	
AUTHORS	Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.	/organism="Hepatitis C virus"	
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use	/mol_type="genomic DNA"	
JOURNAL	Patent: US 6245503-A 23 12-JUN-2001;	/db_xref="taxon:11103"	
FEATURES	Location/Qualifiers	1..558	
source	1..561	/note="unnamed protein product"	
BASE COUNT	103 a 176 c 155 g 127 t	/codon_start=1	
ORIGIN	/organism="unknown"	/protein_id="CAD38073.1"	
Query Match 87.4%; Score 556; DB 6; Length 561;			
Best Local Similarity 100.0%; Pred. No. 2.7e-112;			
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGTTGGGTAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
QY	61	CGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGCGCTCGCGCATGGCTCCGGGT	120
Db	61	CGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGCGCTCGCGCATGGCTCCGGGT	120
QY	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG	240
QY	241	TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
Db	241	TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
QY	301	GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAAACAATCTTCCCGC	360
Db	301	GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAAACAATCTTCCCGC	360
QY	361	TGCTGGTAGCGCTCACCCCGCATCGCTAGAGAACCCAGCGTCCCAACGACGACA	420
Db	361	TGCTGGTAGCGCTCACCCCGCATCGCTAGAGAACCCAGCGTCCCAACGACGACA	420
QY	421	ATACGACGACCGCTCGATTCCCGAGCTTTCACCATCTCGCCCTCGCGGATGAGACGGTG	480
Db	421	ATACGACGACCGCTCGATTCCCGAGCTTTCACCATCTCGCCCTCGCGGATGAGACGGTG	480
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGCTTGGGAT	540
Db	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGCTTGGGAT	540
QY	541	ATGATGATGAACCTGGT	556
Db	541	ATGATGATGAACCTGGT	556
RESULT 15			
AX452772			
LOCUS	561 bp	DNA	linear
DEFINITION	Sequence 23 from Patent EP1211315.		
ACCESSION	AX452772		
VERSION	AX452772.1 GI:21712457		
KEYWORDS	Hepatitis C virus		
SOURCE	Hepatitis C virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1		
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.		
TITLE	Recombinant vectors for producing hcv envelope proteins		
JOURNAL	Patent: EP 1211315-A 23 05-JUN-2002; Innogenetics N.V. (BE)		

FEATURES		Location/Qualifiers	
source	1..561	/organism="Hepatitis C virus"	
CDS	1..558	/mol_type="genomic DNA"	
	1..558	/db_xref="taxon:11103"	
	1..558	/note="unnamed protein product"	
	1..555	/codon_start=1	
	1..555	/protein_id="CAD38073.1"	
	1..555	/db_xref="GI:21712458"	
	1..555	/translation="MLGKVIDTLTCGFDLVGVIPLVGAPLGGAARALAHGVRVLBDG VNVATGNLPGCSFSIFLLALLSCLTPASAYEVNRVSYMHVNTNDCSNSSIVVEAADM IMHTPGCVPCVRENNSRRCVVALTPTLAARNASVPTTIRRHVDSQLFTISPRHETV QDCNSIYPGHITGHRMAWDMMNW"	
mat_peptide	1..555	/product="unnamed"	
BASE COUNT	103 a 176 c 155 g 127 t		
ORIGIN			
Query Match 87.4%; Score 556; DB 6; Length 561;			
Best Local Similarity 100.0%; Pred. No. 2.7e-112;			
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGTTGGGTAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATT	60
QY	61	CGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGCGCTTCGCCGATGGCTCCGGGT	120
Db	61	CGCTCGTGGCGCCCCCTAGGGGCGCTGCCAGGCGCTTCGCCGATGGCTCCGGGT	120
QY	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTATCTTC	180
Db	121	CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTATCTTC	180
QY	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTG	240
QY	241	TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
Db	241	TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG	300
QY	301	GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAAACAATCTTCCCGC	360
Db	301	GACATGATCATGCACACCCCGGGTGGTGGCTTCGGGAGAAACAATCTTCCCGC	360
QY	361	TGCTGGTAGCGCTCACCCCGCATCGCTAGAGAACCCAGCGTCCCAACGACGACA	420
Db	361	TGCTGGTAGCGCTCACCCCGCATCGCTAGAGAACCCAGCGTCCCAACGACGACA	420
QY	421	ATACGACGACCGCTCGATTCCCGAGCTTTCACCATCTCGCCCTCGCGGATGAGACGGTG	480
Db	421	ATACGACGACCGCTCGATTCCCGAGCTTTCACCATCTCGCCCTCGCGGATGAGACGGTG	480
QY	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGCTTGGGAT	540
Db	481	CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGCTTGGGAT	540
QY	541	ATGATGATGAACCTGGT	556
Db	541	ATGATGATGAACCTGGT	556
RESULT 15			
AX452772			
LOCUS	561 bp	DNA	linear
DEFINITION	Sequence 23 from Patent EP1211315.		
ACCESSION	AX452772		
VERSION	AX452772.1 GI:21712457		
KEYWORDS	Hepatitis C virus		
SOURCE	Hepatitis C virus		
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1		
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.		
TITLE	Recombinant vectors for producing hcv envelope proteins		
JOURNAL	Patent: EP 1211315-A 23 05-JUN-2002; Innogenetics N.V. (BE)		

Search completed: December 20, 2003, 02:02:05
Job time : 2410.07 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 175.828 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303a-27

Perfect score: 636

Sequence: 1 ATGTTGGTAAGGTCATCGA.....ATCACCACCATCACTAATAG 636

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	636	100.0	636	17	AAT12964 HCV E1 construct H
2	636	100.0	636	24	AAL48928 Hepatitis C virus
3	597.4	93.9	606	17	AAT12963 HCV E1 construct H
4	597.4	93.9	606	24	AAL48927 Hepatitis C virus
5	597	93.9	723	17	AAT12961 HCV E1 construct H
6	597	93.9	723	24	AAL48925 Hepatitis C virus
7	556	87.4	561	17	AAT12962 HCV E1 construct H
8	556	87.4	561	24	AAL48926 Hepatitis C virus

9	515	81.0	795	17	AAT12705 HCV E1 construct H
10	515	81.0	795	24	AAL48914 Hepatitis C virus
11	512	80.5	2082	24	AAL48939 Hepatitis C virus
12	512	80.5	2086	17	AAT12973 HCV E1 construct H
13	512	80.5	2433	17	AAT12974 HCV E1 construct H
14	501	78.8	2434	24	AAL48940 Hepatitis C virus
15	456.4	71.8	633	17	AAT12706 HCV E1 construct H
16	456.4	71.8	633	24	AAL48915 Hepatitis C virus
17	451.6	71.0	636	17	AAT12709 HCV E1 construct H
18	451.6	71.0	636	24	AAL48918 Hepatitis C virus
19	451.4	71.0	673	19	AAV42305 Cuticle protein c
20	451.2	70.9	2187	19	ABA03491 NANBV hepatitis vir
21	451.2	70.9	2540	14	AAQ43889 Hepatitis C virus
22	451.2	70.9	2540	15	AAQ63753 Hepatitis C virus
23	451.2	70.9	9605	24	ABK91411 Hepatitis C virus
24	451.2	70.9	9605	24	ABK91424 Hepatitis C virus
25	451.2	70.9	9605	24	ABK91425 Hepatitis C virus
26	451.2	70.9	9605	24	ABK91426 Hepatitis C virus
27	451.2	70.9	9605	24	ABK91428 Hepatitis C virus
28	451.2	70.9	9605	24	ABK91429 Hepatitis C virus
29	451.2	70.9	9605	24	ABK91430 Hepatitis C virus
30	451.2	70.9	9605	24	ABK91431 Hepatitis C virus
31	451.2	70.9	9605	24	ABK91432 Hepatitis C virus
32	451.2	70.9	9605	24	ABK91433 Hepatitis C virus
33	451.2	70.9	9605	24	AAQ25332 Hepatitis C virus
34	451.2	70.9	9608	24	ABK91427 Hepatitis C virus
35	451.2	70.9	11062	24	AAQ25331 Hepatitis C virus
36	451.2	70.9	11076	21	AAA98965 Hepatitis C virus
37	449.6	70.7	1562	19	AAV60672 Fragment #5 isolat
38	449.6	70.7	1880	13	AAQ24467 NANBV hepatitis vir
39	449.6	70.7	1953	25	AAQ55222 Plasmid pIDK2 DNA
40	449.6	70.7	2829	19	AAV60673 Fragment #6 isolat
41	448	70.4	1251	13	AAQ26981 HCV gene 1. Hepat
42	448	70.4	2540	13	AAQ29628 Hepatitis C virus
43	448	70.4	3360	17	AAT03677 Hepatitis C genome
44	448	70.4	3461	15	AAQ64068 Non-A, non-B hepat
45	448	70.4	3461	16	AAT30386 5'UTR/CORE/ENV/NS1

ALIGNMENTS

RESULT 1
AAT12964
ID AAT12964 standard; DNA; 636 BP.

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

XX AC AAT12964;

Qy

181 CTTGGGCTTTGCTGTCCCTGTCTGACCGTCCAGCTTCCGCTTATGAAGTGC

181 CTCTTGGCTTTGCTGTCCCTGTCTGACCGTTCAGCTTCCGCTTATGAAGTGCGCAACGTG 240


```
Best Local Similarity 100.0%; Pred. No. 3.5e-148;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTATCATGATACCCCTTACATCGGGCTTCGCCAGCCTCTGGGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCATGATACCCCTTACATCGGGCTTCGCCAGCCTCTGGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGATGGCTCCGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGATGGCTCCGGTT 120
QY 121 CTGGAGGACGGGCTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGCTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGTCTGTCTGACCGTTCAGCTTCGCTTATGAAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTGCTGTCTGTCTGACCGTTCAGCTTCGCTTATGAAGTGGCAACGTG 240
QY 241 TCGGGATGATCATGTATGCAAGCACTGCTCAAGCTTCGCTTATGAAGTGGCAACGTG 300
DB 241 TCGGGATGATCATGTATGCAAGCACTGCTCAAGCTTCGCTTATGAAGTGGCAACGTG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCGCCCTGGCGATGGCTCCGGTT 360
DB 301 GACATGATCATGCACACCCCGGGTGGTGGCGCCCTGGCGATGGCTCCGGTT 360
QY 361 TCGTGGGTAGCGCTCACCCCGCGCTCGAGTGGAGAACCCAGCGTCCCCACGACA 420
DB 361 TCGTGGGTAGCGCTCACCCCGCGCTCGAGTGGAGAACCCAGCGTCCCCACGACA 420
QY 421 ATACGACGCCAGCTCGATTCCCGAGCTGTTCAACATCTCGCCCTCGCGGATGAGACGGTG 480
DB 421 ATACGACGCCAGCTCGATTCCCGAGCTGTTCAACATCTCGCCCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
QY 541 ATGATGATGAACCTGGTTCGCTTACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATC 597
DB 541 ATGATGATGAACCTGGTTCGCTTACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATC 597

RESULT 6
AAL48925
ID AAL48925 standard; DNA; 723 BP.
XX
AC AAL48925;
XX
DE Hepatitis C virus clone HC137 E1 protein coding sequence.
XX
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
XX virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
PN WO20025548-A2.
XX
PD 18-JUL-2002.
XX
PF 11-JAN-2002; 2002WO-BP00219.
XX
PR 11-JAN-2001; 2001US-260699P.
XX
PR 30-AUG-2001; 2001US-315768P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Maertens G, Bosman F, Buyse M;
XX
DR WPI; 2002-599657/64.
```

```
P-PSDB; AAO18666.
New therapeutic vaccine compositions comprising at least one purified
recombinant hepatitis C virus (HCV) single or specific oligomeric
recombinant envelope protein E1 or E2, useful for immunizing humans
from HCV infection
Example 2; Page 173-174; 243pp; English.
The present invention relates to new therapeutic vaccine compositions for
inducing hepatitis C virus (HCV)-specific antibodies, comprising a
composition containing at least one purified recombinant HCV single or
specific oligomeric recombinant envelope proteins selected from an E1 and
an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
useful for inducing HCV-specific antibodies or for immunizing humans
against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
vaccines or therapeutics, in HCV screening and confirmatory antibody
tests, for raising antibodies, in the preparation of medicament, and for
in vitro monitoring of HCV disease or prognosing the response to
treatment of patients suffering from HCV infection. The present sequence
is a coding sequence described in the exemplification of the invention.
Sequence 723 BP; 126 A; 220 C; 208 G; 169 T; 0 other;
Query Match 93.9%; Score 597; DB 24; Length 723;
Best Local Similarity 100.0%; Pred. No. 3.5e-148;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGGTAAGGTATCATGATACCCCTTACATCGGGCTTCGCCAGCCTCTGGGGGTACATT 60
DB 1 ATGTTGGGTAAGGTATCATGATACCCCTTACATCGGGCTTCGCCAGCCTCTGGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGATGGCTCCGGTT 120
DB 61 CCGCTCGTCGGCGCCCCCTAGGGGCGCTGCCAGGGCCCTGGCGATGGCTCCGGTT 120
QY 121 CTGGAGGACGGGCTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGCTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
QY 181 CTCCTGGCTTTGCTGTCTGTCTGACCGTTCAGCTTCGCTTATGAAGTGGCAACGTG 240
DB 181 CTCCTGGCTTTGCTGTCTGTCTGACCGTTCAGCTTCGCTTATGAAGTGGCAACGTG 240
QY 241 TCGGGATGATCATGCACACCCCGGGTGGTGGCGCCCTGGCGATGGCTCCGGTT 300
DB 241 TCGGGATGATCATGCACACCCCGGGTGGTGGCGCCCTGGCGATGGCTCCGGTT 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCGCCCTGGCGATGGCTCCGGTT 360
DB 301 GACATGATCATGCACACCCCGGGTGGTGGCGCCCTGGCGATGGCTCCGGTT 360
QY 361 TCGTGGGTAGCGCTCACCCCGCGCTCGAGTGGAGAACCCAGCGTCCCCACGACA 420
DB 361 TCGTGGGTAGCGCTCACCCCGCGCTCGAGTGGAGAACCCAGCGTCCCCACGACA 420
QY 421 ATACGACGCCAGCTCGATTCCCGAGCTGTTCAACATCTCGCCCTCGCGGATGAGACGGTG 480
DB 421 ATACGACGCCAGCTCGATTCCCGAGCTGTTCAACATCTCGCCCTCGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCAACGATGGCTGGGAT 540
QY 541 ATGATGATGAACCTGGTTCGCTTACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATC 597
DB 541 ATGATGATGAACCTGGTTCGCTTACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATC 597

RESULT 7
AAT12962
ID AAT12962 standard; DNA; 561 BP.
XX
```

AC	AAT12962;	QY	241	TCCGGGATGTACCATGTCAAGAAAGAGTGTCCCAACTCAAGCATTTGTATGAGCAGCG	300
XX					
DT	24-SEP-1996 (first entry)	Db	241	TCCGGGATGTACCATGTCAAGAAAGAGTGTCCCAACTCAAGCATTTGTATGAGCAGCG	300
XX					
DE	HCV E1 construct HCC138.				
XX					
XX	HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;	QY	301	GACATGATCATGACACACCCCGGGTGGCTGCGCTTCGGGTTCGGGAGAACAACTCTTCCGCG	360
XX					
KW	serotype; reversed phase hybridisation assay; genotype; antigen; sera;	Db	301	GACATGATCATGACACACCCCGGGTGGCTGCGCTTCGGGTTCGGGAGAACAACTCTTCCGCG	360
XX	ss.				
XX		QY	361	TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCCAGCGTCCCAACACGACA	420
OS	Hepatitis C virus.	Db	361	TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCCAGCGTCCCAACACGACA	420
XX					
PN	WO9604385-A2.	QY	421	ATACGACGCCACGTCGATTCGCCAGCTGTTACCATCTCGCCTCGCGGATGAGACGGTG	480
XX					
XX	15-FEB-1996.	Db	421	ATACGACGCCACGTCGATTCGCCAGCTGTTACCATCTCGCCTCGCGGATGAGACGGTG	480
XX					
XX	31-JUL-1995; 95WO-BP03031.	QY	481	CAGGACTGCAATTCGCTCAATCTATCCCGGCCACATACCGGGTCAACGATGCGTTGGGAT	540
XX					
XX	29-JUL-1994; 94EP-0870132.	Db	481	CAGGACTGCAATTCGCTCAATCTATCCCGGCCACATACCGGGTCAACGATGCGTTGGGAT	540
XX	(INNO-) INNOGENETICS NV.				
XX		QY	541	ATGATGATGAACCTGGT 556	
XX	Bosman F, Buyse M, De Martynoff G, Maertens G;				
XX	WPI; 1996-129401/13.	Db	541	ATGATGATGAACCTGGT 556	
DR					
XX	Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope				
XX	proteins - in presence of disulphide bond cleavage agent, to				
PT	produce proteins suitable for direct use in vaccines of diagnostic				
PT	assays of HCV				
XX					
XX	Claim 23; Fig 21; 146pp; English.				
PS					
XX	AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1				
XX	and E2 protein coding sequence constructs. These sequences are included				
CC	in vectors for the production of recombinant E1, E2, and E1/E2 proteins.				
CC	The recombinant proteins can then be isolated using a method of the				
CC	invention. In the method, the envelope proteins are purified by				
CC	carrying out a disulphide bond cleavage, or a reduction step with a				
CC	disulphide bond cleavage agent, after lysis of recombinant host cells.				
CC	The constructs containing the purified HCV envelope proteins can be used				
CC	for vaccinating humans against HCV, for in vitro detection of HCV				
CC	antibodies in a sample, and in a serotyping assay for detecting one or				
CC	more serological types of HCV present in a biological sample. The				
CC	constructs can also be immobilised on a solid substrate and incorporated				
CC	into a reversed phase hybridisation assay for determining the presence or				
CC	the genotype of HCV. The new purification method preserves the				
CC	conformation of the recombinantly expressed E1, E2 and E1/E2, and				
CC	eliminates contaminating proteins. Antigens isolated using this method				
CC	are more reactive with human sera than those isolated by known				
XX	techniques.				
XX					
SQ	Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;				
	Query Match				
	Best Local Similarity 87.4%; Score 556; DB 17; Length 561;				
	Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 ATGTTGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT	60			
Db	1 ATGTTGGTAAAGTATCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATT	60			
QY	61 CCGCTCGTGGCGCCCGCTAGGGGCGCTGCAGGGCGCTGGCGATGGCGTTCGGGTT	120			
Db	61 CCGCTCGTGGCGCCCGCTAGGGGCGCTGCAGGGCGCTGGCGATGGCGTTCGGGTT	120			
QY	121 CTGAGGACGGGTGAATATGACAGGGCAATTTGCCGGTGTCTTTCTTCTATCTTC	180			
Db	121 CTGAGGACGGGTGAATATGACAGGGCAATTTGCCGGTGTCTTTCTTCTATCTTC	180			
QY	181 CTCTTGGCTTTGCTGCTGCTGTGACAGCTTCAGCTTCAGCTATGAGTGGCAACGTG	240			
Db	181 CTCTTGGCTTTGCTGCTGCTGTGACAGCTTCAGCTTCAGCTATGAGTGGCAACGTG	240			

Example 2; Page 175-176; 243pp; English.

The present invention relates to new therapeutic vaccine compositions for inducing hepatitis C virus (HCV)-specific antibodies, comprising a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from an E1 and an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are useful for inducing HCV-specific antibodies or for immunising humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for

CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 561 BP; 103 A; 176 C; 155 G; 127 T; 0 other;

Query Match 87.4%; Score 556; DB 24; Length 561;
Best Local Similarity 100.0%; Pred. No. 2.3e-137;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTGGTAAAGTTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGTAAAGTTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
Qy 61 CCGCTGTCGGGGCCCCCTAGGGGGCGCTGCGAGGCGCTGCGCATGGGTCGCGGT 120
Db 61 CCGCTGTCGGGGCCCCCTAGGGGGCGCTGCGAGGCGCTGCGCATGGGTCGCGGT 120
Qy 121 CTGGAGGACGGCTGAATATGCAACAGGAAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAATATGCAACAGGAAATTTGCCGGTGTCTTCTCTATCTTC 180
Qy 181 CTCCTGGCTTTGCTGCTCTGCTGACCGCTTCAGCTTCGCGTTATGAAGTGCACACGTG 240
Db 181 CTCCTGGCTTTGCTGCTGCTGACCGCTTCAGCTTCGCGTTATGAAGTGCACACGTG 240
Qy 241 TCCGGGATGATACCTGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
Db 241 TCCGGGATGATACCTGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
Qy 301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACATCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACATCTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTCAGCCCGGCGCTGCGAGTACGAGCGGCGGCGGCGGCGGCGG 420
Db 361 TGCTGGGTAGCGCTCAGCCCGGCGCTGCGAGTACGAGCGGCGGCGGCGGCGGCGG 420
Qy 421 ATAGTATGATGAAGTGGT 556
Db 421 ATAGTATGATGAAGTGGT 556

RESULT 9

AATL12705
ID AATL12705 standard; DNA; 795 BP.
XX
AC AATL12705;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC110A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW 88.
XX
OS Hepatitis C virus.
XX
PN WO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX

PR 29-JUL-1994; 94EP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
PI Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX
XX AATL12704-T12709 and AATL12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
XX techniques.
SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 81.0%; Score 515; DB 17; Length 795;
Best Local Similarity 89.2%; Pred. No. 1.8e-126;
Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
Qy 1 ATGTTGGTAAAGTTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
Db 1 ATGTTGGTAAAGTTCATGATACCTTACATGCGGCTTCGCGACCTCGTGGGTACATT 60
Qy 61 CCGCTGCTCGGCGCCCCCTAGGGGGCGCTGCGAGGCGCTTGGCGCATGGCGGT 120
Db 61 CCGCTGCTCGGCGCCCCCTAGGGGGCGCTGCGAGGCGCTTGGCGCATGGCGGT 120
Qy 121 CTGGAGGACGGCTGAATATGCAACAGGAAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGCTGAATATGCAACAGGAAATTTGCCGGTGTCTTCTCTATCTTC 180
Qy 181 CTCCTGGCTTTGCTGCTGCTGCTGACCGCTTCAGCTTCGCGTTATGAAGTGCACACGTG 240
Db 181 CTCCTGGCTTTGCTGCTGCTGCTGACCGCTTCAGCTTCGCGTTATGAAGTGCACACGTG 240
Qy 241 TCCGGGATGATACCTGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
Db 241 TCCGGGATGATACCTGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
Qy 301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACATCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACATCTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTCAGCCCGGCGCTGCGAGTACGAGCGGCGGCGGCGGCGGCGG 420
Db 361 TGCTGGGTAGCGCTCAGCCCGGCGCTGCGAGTACGAGCGGCGGCGGCGGCGGCGG 420
Qy 421 ATAGTATGATGAAGTGGT-----
Db 421 ATAGTATGATGAAGTGGT-----
Qy 439 -----TCCAGCTGTTTCCACCATCTCGCGCTCGCGCG 468

1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT	60
1	ATGTTGGGTAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT	60
61	CGCTCGTCGGCGGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
61	CGCTCGTCGGCGGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120
121	CTGGAGACGGCGGTGAACCTATCAACAGGGAAATTTGCCGGTTGCTCTTTCTCTATCTTC	180
121	CTGGAGACGGCGGTGAACCTATCAACAGGGAAATTTGCCGGTTGCTCTTTCTCTATCTTC	180
181	CTCTTGCTTTGCTGTCCTGTGACCGTTCCAGCTTCGGCTTATGAAGTCGCAACGTG	240
181	CTCTTGCTTTGCTGTCCTGTGACCGTTTCAGCTTTCGCTTATGAAGTCGCAACGTG	240
241	TCGGGATGTACCATGTCACGAACGACTCTCAAATCAAGATTGTGTATGAGCAGCG	300
241	TCGGGATGTACCATGTCACGAACGACTCTCAAATCAAGATTGTGTATGAGCAGCG	300
301	GACATGATCATGACACCCCGGGTGGCTCGCTGTTCCGGAGACAACTCTTCCCGC	360
301	GACATGATCATGACACCCCGGGTGGCTCGCTGTTCCGGAGACAACTCTTCCCGC	360
361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAAGCCAGCGTCCCACACGACA	420
361	TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAAAGCCAGCGTCCCACACGACA	420
421	ATACGACGCCACGTCGAT-----	438
421	ATACGACGCCACGTCGATTTGCTCGTTGGGGCGCTGCTTCTGTTCGCTATGTACGTG	480
439	-----TCCGACGCTTACCAATCTCGCTCGCGCG	468
481	GGGGACCTCTGGGATCTGTTCTTCTCGTCTCCAGCTGTTCCACCATCGCCTCCCGG	540
469	CATGACGGTGCAGGACTGCAAATGCTCAATCTATCCCGGCCACACATAACGGGTCAACGT	528
541	CATGACGGTGCAGGACTGCAAATGCTCAATCTATCCCGGCCACACATAACGGGTCAACGT	600
529	ATGGCTTGGGATATGATGAACTGGTTCGCTTCAAACGGCCCTGGTGGTATCGCAGCTG	588
601	ATGGCTTGGGATATGATGAACTGGTTCGCTTCAAACGGCCCTGGTGGTATCGCAGCTG	660
589	CTCCGGATC	597
661	CTCCGGATC	669

RESULT 11	
AA148939	
ID	AA148939 standard; DNA; 2082 BP.
XX	
XX	
AC	AA148939;
XX	
DT	24-OCT-2002 (first entry)
XX	
XX	Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.
DE	
XX	
XX	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	virucide; immunostimulant; vaccine; ds.
KW	
XX	
OS	Hepatitis C virus.
XX	
PN	WO200255548-A2.
XX	
PD	18-JUL-2002.
XX	
PF	11-JAN-2002; 2002WO-EP00219.
XX	
PR	11-JAN-2001; 2001US-260699P.
PR	30-AUG-2001; 2001US-315768P.
XX	

b	481	GGGGACCTCTGGGATCTGCTCTTCCTCGTCTCCAGACTGTTACCAATCTCGCCTCGCCGG	540
y	469	CATGAGACGGTGACGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTCAACGCT	528
b	541	CATGAGACGGTGACGACTGCAATTGCTCAATCTATCCGGGCCACATAACGGGTCAACGCT	600
y	529	ATGGCTTGGGATATGATGATGAACCTGGTCGCCCTACAACGGCCCTGGTGATCGCAGCTG	588
b	601	ATGGCTTGGGATATGATGATGAACCTGGTCGCCCTACAACGGCCCTGGTGATCGCAGCTG	660
y	589	CTCCGGATC 597	
b	661	CTCCGGATC 669	
	RESULT 10		
	AAI48914		
	D	AAI48914 standard; DNA; 795 BP.	
X	X	AAI48914;	
X	X	24-OCT-2002 (first entry)	
X	X	Hepatitis C virus clone HCV110A E1 protein coding sequence.	
X	X	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;	
X	X	virucide; immunostimulant; vaccine; ds.	
X	X	Hepatitis C virus.	
DS	XX		
XX	XX	WO200255548-A2.	
XX	XX	18-JUL-2002.	
XX	XX	11-JAN-2002; 2002WO-EP00219.	
XX	XX	11-JAN-2001; 2001US-260699P.	
XX	XX	30-AUG-2001; 2001US-315768P.	
XX	XX	(INNO-) INNOGENETICS NV.	
XX	XX	Maertens G, Bosman F, Buyse M;	
XX	XX	WPI; 2002-599657/64.	
DR	DR	P-PSDB; AAO18661.	
XX	XX	New therapeutic vaccine compositions comprising at least one purified	
PT	PT	recombinant hepatitis C virus (HCV) single or specific oligomeric	
PT	PT	recombinant envelope protein E1 or E2, useful for immunizing humans	
PT	PT	from HCV infection	
XX	XX	Example 2; Page 161-162; 243pp; English.	
XX	XX	The present invention relates to new therapeutic vaccine compositions for	
CC	CC	inducing hepatitis C virus (HCV)-specific antibodies, comprising a	
CC	CC	composition containing at least one purified recombinant HCV single or	
CC	CC	specific oligomeric recombinant envelope protein selected from an E1 and	
CC	CC	an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are	
CC	CC	useful for inducing HCV-specific antibodies or for immunising humans	
CC	CC	against HCV. The recombinant HCV E1 and/or E2 proteins are useful as	
CC	CC	vaccines or therapeutics, in HCV screening and confirmatory antibody	
CC	CC	tests, for raising antibodies, in the preparation of medicament, and for	
CC	CC	in vitro monitoring of HCV disease or prognosing the response to	
CC	CC	treatment of patients suffering from HCV infection. The present sequence	
CC	CC	is a coding sequence described in the exemplification of the invention.	
XX	XX	Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;	
XX	XX		
XX	XX	Query Match 81.0%; Score 515; DB 24; Length 795;	
XX	XX	Best Local Similarity 89.2%; Pred. No. 1.8e-126;	
XX	XX	Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;	

```
PA (INNO-) INNOGENETICS NV.
XX
PI Maertens G, Bosman F, Buyse M;
XX
DR WPI; 2002-599657/64.
DR P-PSDB; AAO18678.
XX
PT New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans
PT from HCV infection
XX
XX Example 2; Page 206-209; 243pp; English.
XX
CC The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention.
XX
SQ Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 other;

Query Match
Best Local Similarity 80.5%; Score 512; DB 24; Length 2082;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGTAAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
DB 4 TTGGTAAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63

QY 64 CTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
DB 64 CTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123

QY 124 GAGGACGGGTGAATATGCAACAGGAATTTGCCGGTGTCTTCTATCTTCCTC 183
DB 124 GAGGACGGGTGAATATGCAACAGGAATTTGCCGGTGTCTTCTATCTTCCTC 183

QY 184 TTGGCTTTGCTCTCTGTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTGCC 243
DB 184 TTGGCTTTGCTCTCTGTCTGACCGTTCCAGCTTCGGCTTATGAAGTGGCAACGTGCC 243

QY 244 GGGATGTACCATGTCCAGAACCACTGCTCCAACTCAAGCATTTGTGTATGAGGACGGAC 303
DB 244 GGGATGTACCATGTCCAGAACCACTGCTCCAACTCAAGCATTTGTGTATGAGGACGGAC 303

QY 304 ATGATCATGCACACCCCGGTTGCTGCGCTCGGTTCCGGAGAACACTTCTCCCGTGC 363
DB 304 ATGATCATGCACACCCCGGTTGCTGCGCTCGGTTCCGGAGAACACTTCTCCCGTGC 363

QY 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAGCCAGCGTCCCAACACGACAATA 423
DB 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAGCCAGCGTCCCAACACGACAATA 423

QY 424 CGACGCCACGTCGAT-----TCCAGCTGTTTCCGCTATGATGACGTGGG 483
DB 424 CGACGCCACGTCGATTTGCTCGTTGGGGGGGCTGCTTCTGTTCCGCTATGATGACGTGGG 483

QY 439 -----TCCAGCTGTTTCCGCTATGATGACGTGGG 471
DB 484 GACCTCTGGGATCTGTCTTCTCGTCTCCAGCTGTTTCCAGCTATGATGACGTGGG 543

QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATACGGGTACCGTATG 531
DB 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATACGGGTACCGTATG 603
```

PA 532 GCTTGGGATATGATGATGAACCTGGTTCGCTCAACAGGCCCTGGTGGTATCGCAGCTGCTC 591
DB 604 GCTTGGGATATGATGATGAACCTGGTTCGCTCAACAGGCCCTGGTGGTATCGCAGCTGCTC 663
QY 592 CGGATC 597
DB 664 CGGATC 669
RESULT 12
AAT12973
ID AAT12973 standard; DNA; 2086 BP.
XX AAT12973;
AC AAT12973;
DT 24-SEP-1996 (first entry)
XX HCV E1 construct HCC165.
DB HCV E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX Hepatitis C virus.
OS Hepatitis C virus.
XX WO9604385-A2.
PN 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
PF 29-JUL-1994; 94EP-0870132.
PR (INNO-) INNOGENETICS NV.
PA Bosman F, Buyse M, De Martynoff G, Maertens G;
PI WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX Claim 23; Fig 21; 146pp; English.
PS AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;
SQ Query Match 80.5%; Score 512; DB 17; Length 2086;
Best Local Similarity 89.2%; Pred. No. 1,4e-125;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 4 TTGGTAAAGGTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
|||||

Db 4 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTTCCG 63
QY 64 CTCGTGCGCGCCCTTAGGGGGCGCTGCAGGGCCCTGGGGCATGGGTCCGGTTCTG 123
Db 64 CTCGTGCGCGCCCTTAGGGGGCGCTGCAGGGCCCTGGGGCATGGGTCCGGTTCTG 123
QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTCTC 183
Db 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTCTC 183
QY 184 TTGGCTTTGCTCTCTGCTGCTGACCGTTCCAGTTCCGCTTATGAAGTGCAGCGTTC 243
Db 184 TTGGCTTTGCTCTCTGCTGCTGACCGTTCCAGTTCCGCTTATGAAGTGCAGCGTTC 243
QY 244 GGGATGTACATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGGAC 303
Db 244 GGGATGTACATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGGAC 303
QY 304 ATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTTTCCCGTGC 363
Db 304 ATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTTTCCCGTGC 363
QY 364 TGGGTAGCGCTCACCCCGGCTCGAGTGGAGACCGCGTCCCAACAGCAATA 423
Db 364 TGGGTAGCGCTCACCCCGGCTCGAGTGGAGACCGCGTCCCAACAGCAATA 423
QY 424 CGACGCCACGTGCAT----- 438
Db 424 CGACGCCACGTGCATTTGCTGTGGGGGGGTGCTTTCTGTTCCGCTATGTACGTGGG 483
QY 439 -----TCCAGCTTTCAACATCTCGCTCGCGGAT 471
Db 484 GACCTCTGCGGATCTGCTTCTCGTCTCCAGCTGTTCAACATCTCGCTCGCGGAT 543
QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCAATAACGGGTACCGTATG 531
Db 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCAATAACGGGTACCGTATG 603
QY 532 GCTTGGGATATGATGATGAATGCTGCTGCTTACCAACCGCCCTGTTGATGCGAGCTGCTC 591
Db 604 GCTTGGGATATGATGATGAATGCTGCTGCTTACCAACCGCCCTGTTGATGCGAGCTGCTC 663
QY 592 CGGATC 597
Db 664 CGGATC 669

RESULT 13
AAT12974
ID AAT12974 standard; DNA; 2433 BP.
XX AC AAT12974;
XX DT 25-SEP-1996 (first entry)
XX HCV E1 construct HCC166.
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX Hepatitis C virus.
XX WO9604385-A2.
XX 15-FEB-1996.
XX 31-JUL-1995; 95WO-EP03031.
XX 29-JUL-1994; 94EP-0870132.
XX (INNO-) INNOGENETICS NV.
XX

PI XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of disulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
XX assays of HCV
XX Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.

XX Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;
SQ
Query Match 80.5%; Score 512; DB 17; Length 2433;
Best Local Similarity 89.2%; Pred. No. 1.5e-125;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
QY 4 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTTCCG 63
Db 355 TTGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTTCCG 414
QY 64 CTCGTGCGCGCCCTTAGGGGGCGCTGCAGGGCCCTGGGGCATGGGTCCGGTTCTG 123
Db 415 CTCGTGCGCGCCCTTAGGGGGCGCTGCAGGGCCCTGGGGCATGGGTCCGGTTCTG 474
QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTCTC 183
Db 475 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTCCTTCTCTATCTTCTC 534
QY 184 TTGGCTTTGCTCTCTGCTGCTGACCGTTCCAGTTCCGCTTATGAAGTGCAGCGTTC 243
Db 535 TTGGCTTTGCTCTCTGCTGCTGACCGTTCCAGTTCCGCTTATGAAGTGCAGCGTTC 594
QY 244 GGGATGTACATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGGAC 303
Db 595 GGGATGTACATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGGAC 654
QY 304 ATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTTTCCCGTGC 363
Db 655 ATGATCATGCACACCCCGGGTGGTGCCTCGCTTCGGGAGAACAACTTTCCCGTGC 714
QY 364 TGGGTAGCGCTCACCCCGGCTCGAGTGGAGACCGCGTCCCAACAGCAATA 423
Db 715 TGGGTAGCGCTCACCCCGGCTCGAGTGGAGACCGCGTCCCAACAGCAATA 774
QY 424 CGACGCCACGTGCAT----- 438
Db 775 CGACGCCACGTGCATTTGCTGTGGGGGGGTGCTTTCTGTTCCGCTATGTACGTGGG 834
QY 439 -----TCCAGCTTTCAACATCTCGCTCGCGGAT 471
Db 835 GACCTCTGCGGATCTGCTTCTCGTCTCCAGCTGTTCAACATCTCGCTCGCGGAT 894
QY 472 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCAATAACGGGTACCGGTATG 531


```
Db      895 GAGCGGTGCGAGCTGCAATTGCTCAATCTATCCCGGCCACATACGCGGTACCGGTATG 954
Qy      532 GCTTGGGATATGATGATGAACCTGGTCCGCTCAACAGCGCCCTGGTGTATGCGAGCTGTCTC 591
Db      955 GCTTGGGATATGATGATGAACCTGGTCCGCTCAACAGCGCCCTGGTGTATGCGAGCTGTCTC 1014
Qy      592 CGGATC 597
Db      1015 CGGATC 1020

RESULT 14
ID AAL48940 standard; DNA; 2434 BP.
AC AAL48940;
DT 24-OCT-2002 (first entry)
XX
DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX
OS Hepatitis C virus.
XX
XX WO20025548-A2.
XX
XX 18-JUL-2002.
XX
XX 11-JAN-2002; 2002WO-EP00219.
XX
XX 11-JAN-2001; 2001US-260699P.
XX 30-AUG-2001; 2001US-315768P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Bosman F, Buyse M;
XX
XX WPI; 2002-599657/64.
XX P-PSDB; AAO18679.
XX
XX New therapeutic vaccine compositions comprising at least one purified
XX recombinant hepatitis C virus (HCV) single or specific oligomeric
XX recombinant envelope protein E1 or E2, useful for immunizing humans
XX from HCV infection.
XX
XX Example 2; Page 212-215; 243pp; English.
XX
XX The present invention relates to new therapeutic vaccine compositions for
XX inducing hepatitis C virus (HCV)-specific antibodies, comprising a
XX composition containing at least one purified recombinant HCV single or
XX specific oligomeric recombinant envelope proteins selected from an E1 and
XX an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
XX useful for inducing HCV-specific antibodies or for immunising humans
XX against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
XX vaccines or therapeutics, in HCV screening and confirmatory antibody
XX tests, for raising antibodies, in the preparation of medicament, and for
XX in vitro monitoring of HCV disease or prognosing the response to
XX treatment of patients suffering from HCV infection. The present sequence
XX is a coding sequence described in the exemplification of the invention.
XX
XX Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 other;

Query Match      78.8%; Score 501; DB 24; Length 2434;
Best Local Similarity 89.1%; Pred.No. 1.2e-122;
Matches 594; Conservative 0; Mismatches 0; Indels 73; Gaps 2;

Qy      4 TTGGGTAAAGTATCATGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCG 63
Db      355 TTGGGTAAAGTATCATGATACCCCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCG 414
```

```
Qy      64 CTCGTCCGGCGCCCTAGGGGGCGCTGCGAGGGGCCCTGGCGCATGGGCTCGGGTCTTG 123
Db      415 CTCGTCCGGCGCCCTAGGGGGCGCTGCGAGGGGCCCTGGCGCATGGGCTCGGGTCTTG 474
Qy      124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTTGCTCTTCTCTATCTTCCTC 183
Db      475 GAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTTGCTCTTCTCTATCTTCCTC 534
Qy      184 TTGGCTTTTGCTGCTCCTG-TCTGACCGTTTCCAGCTTCCGCTTATGAAGTGGCAAGGTGTC 242
Db      535 TTGGCTTTTGCTGCTCCTGTTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGGTGTC 594
Qy      243 CGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCGGA 302
Db      595 CGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGAGCGGA 654
Qy      303 CATGATCATGACACCCCGGGTGGTCCCTGCGCTTCCGGGAGAACACTCTTCCCGCTG 362
Db      655 CATGATCATGACACCCCGGGTGGTCCCTGCGCTTCCGGGAGAACACTCTTCCCGCTG 714
Qy      363 CTGGGTAGCGGTCAACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCACACGACAAT 422
Db      715 CTGGGTAGCGGTCAACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCACACGACAAT 774
Qy      423 ACGACGCCACGTGAT-----
Db      775 ACGACGCCACGTGATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTAGTGGG 834
Qy      439 -----TCCAGCTGTTCCACATCTCGCCTCGCGGCA 470
Db      835 GGACCTCTGGGATCTGCTCTCCGCTGTTCCACATCTCGCCTCGCGGCA 894
Qy      471 TGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAAAGGGTCAACGAT 530
Db      895 TGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAAAGGGTCAACGAT 954
Qy      531 GCCTTGGGATATGATGATGAACCTGCTCGCTACAACGGCCCTGGTGGTATCGCAGCTGT 590
Db      955 GCCTTGGGATATGATGATGAACCTGCTCGCTACAACGGCCCTGGTGGTATCGCAGCTGT 1014
Qy      591 CGGATC 597
Db      1015 CGGATC 1021

RESULT 15
AAT12706
ID AAT12706 standard; DNA; 633 BP.
XX
AC AAT12706;
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC11A.
XX
KW HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.
XX
OS Hepatitis C virus.
XX
XX WO9604385-A2.
XX
XX 15-FEB-1996.
XX
XX 31-JUL-1995; 95WO-EP03031.
XX
XX 29-JUL-1994; 94EP-0870132.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
XX
XX
```

Mon Dec 22 13:28:43 2003

QY 529 ATGGCTGGGATATGATGAACTGGT 556
Db 601 ATGGCTGGGATATGATGAACTGGT 628

Search completed: December 19, 2003, 18:51:21
Job time : 178.828 secs

DR WPI; 1996-129401/13.
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
PT proteins - in presence of di-sulphide bond cleavage agent, to
PT produce proteins suitable for direct use in vaccines or diagnostic
PT assays of HCV
XX
PS Claim 23; Fig 21; 146pp; English.
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by a
CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;
Query Match 71.8%; Score 456.4; DB 17; Length 633;
Best Local Similarity 86.8%; Pred. No. 5.3e-111;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;
QY 1 ATGTTGGGTAAGGTCATCGATACCTTACATCGGGCTTGGCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGGTCATCGATACCTTACATCGGGCTTGGCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
Db 61 CCGCTCGTGGGGCCGCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
QY 121 CTGGAGGACGGCGTGAACATGCAACAGGGAAATTGCGCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGCGTGAACATGCAACAGGGAAATTGCGCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTTACGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTTGGCTTTTACGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCCGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCCGGGAGAACAACTCTTCCCGC 360
QY 361 TCGTGGGTAGCGCTCAACCCCGACGCTCGCAGTAGGAACGGCAGCGTCCCACTACGACA 420
Db 361 TCGTGGGTAGCGCTCAACCCCGACGCTCGCAGTAGGAACGGCAGCGTCCCACTACGACA 420
QY 421 ATACGACGCCACGTGGAT----- 438
Db 421 ATACGACGCCACGTGGAT----- 438
QY 439 -----TCCAGCTGTTTCAACCATCTCGCTCGCGG 468
Db 439 -----TCCAGCTGTTTCAACCATCTCGCTCGCGG 468
QY 481 GGGGATCTCTGGGGATCTGTCTCTCTGCTCTCCAGCTTCCAGCATCTCGCTCGCGG 540
Db 481 GGGGATCTCTGGGGATCTGTCTCTCTGCTCTCCAGCTTCCAGCATCTCGCTCGCGG 540
QY 469 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCCCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATACGGGTCCCGT 600

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1636.15 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-27

Perfect score: 636

Sequence: 1 ATGTTGGTAAGGTACGCA.....ATCACCACTCACTAATAG 636

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_estchum:*

3: em_estim:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_nam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	68.2	10.7	488	9 AV755731	AV755731
C 2	54.6	8.6	492	9 AV758366	AV758366
C 3	41.6	6.5	502	12 BI879124	BI879124 fm04e08.y
C 4	40.6	6.4	275	9 AV835132	AV835132

5	40.6	6.4	402	9	AV392783	AV392783
6	40.6	6.4	551	9	AV392165	AV392165
7	40.6	6.4	552	12	BI996341	BI996341
8	40.6	6.4	584	12	BI727879	BI727879
9	40.6	6.4	1201	13	BI356664	BI356664
10	40.2	6.3	1162	12	BM918259	BM918259
11	40.2	6.3	1201	13	BM918259	AGENCOURT
C 12	40.2	6.3	1201	13	BM918259	AGENCOURT
C 13	39	6.1	359	12	BJ252669	BJ252669
C 14	39	6.1	375	12	BJ252669	BJ252669
15	39	6.1	840	29	CC335916	CC335916
16	39	6.1	840	29	CC335916	OGUAI60TV
17	38.6	6.1	925	29	CNS0091P	EL01T0207
C 18	38.4	6.0	606	9	AV915997	AV915997
C 19	38.4	6.0	636	12	BI960110	BI960110
20	38.4	6.0	702	14	CD432549	HVSMEN002
21	38.4	6.0	746	9	AV921112	CD432549 ETH1 30 D
22	38.4	6.0	970	29	CNS010C9	AV921112
C 23	38.4	6.0	987	29	CNS010C9	AV921112
C 24	38.2	6.0	533	6	AU192776	AL098787
C 25	38.2	6.0	538	6	AU193705	AL098787
C 26	38.2	6.0	544	6	AU190971	AL098787
C 27	38.2	6.0	544	6	AU192419	AL098787
C 28	38.2	6.0	1270	12	EG968359	AL098787
C 29	38	6.0	354	14	CB968525	AL098787
30	38	6.0	1201	13	BI381961	AL098787
31	37.8	5.9	435	14	C72860	AL098787
32	37.8	5.9	533	29	CC010084	AL098787
33	37.8	5.9	659	29	CC405164	AL098787
34	37.8	5.9	826	29	BZ736582	AL098787
C 35	37.8	5.9	895	29	CC359028	AL098787
C 36	37.8	5.9	925	29	CC359026	AL098787
C 37	37.8	5.9	940	29	CC010085	AL098787
C 38	37.8	5.9	951	29	CC405167	AL098787
39	37.6	5.9	431	9	AV639153	AL098787
40	37.4	5.9	360	9	AJ473805	AL098787
C 41	37.4	5.9	637	13	BQ293470	AL098787
C 42	37.4	5.9	641	13	BQ172543	AL098787
C 43	37.4	5.9	650	14	CA828039	AL098787
44	37.4	5.9	834	29	BZ641450	AL098787
C 45	37.4	5.9	841	29	BZ641457	AL098787

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

```
source
1..488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAK903"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      116 a 134 c 137 g 97 t
ORIGIN
Query Match      10.7%; Score 68.2; DB 9; Length 488;
Best Local Similarity 72.3%; Pred. No. 4.8e-07;
Matches 102; Conservative 0; Mismatches 36; Indels 1; Gaps 1;

QY 445 CTGTTCCACCATCTCGCTCGCGGATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCG 504
DB 403 CAGCTGATCATCTGGGCTCAGACCATGAGTTTGTGCATGAATGCAACTGCTCCATCTAT 344
QY 505 CCGGGCCACATAACGGGTACCGTATG-CCTTGGGATATGATGATGAACTGTCGCTACAAG 563
DB 343 CCTGGCGCCATCACTGACACCGTATGATGATGGGACATGGGACATGATGAACTGTCGCTGAC 284
QY 564 AACGGCCCTGGTGGTATCGCA 584
DB 283 CGCTGCTATGATCATGCGCGTA 263

RESULT 2
AV758366/c      492 bp      mRNA      linear      EST 19-OCT-2000
LOCUS      AV758366 BM Homo sapiens CDNA clone BMFAKA03 5', mRNA sequence.
DEFINITION      AV758366
ACCESSION      AV758366
VERSION      AV758366.1 GI:10916214
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
,L., Xu,S., Gu,W., Yu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens CDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAKA03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      124 a 128 c 125 g 112 t
ORIGIN
Query Match      8.6%; Score 54.6; DB 9; Length 492;
Best Local Similarity 67.2%; Pred. No. 0.0015;
Matches 92; Conservative 0; Mismatches 44; Indels 1; Gaps 1;
```

```
QY 449 TCACCATCTCGCTCGCGGATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCG 508
DB 400 TGATTATCTCTCAGCAGCAACATTTGTTGTGCAAGATGCACTGCTCTATCTCTG 341
QY 509 GCCACATRAACGGTCT-ACCGTATGCTTGGGATATGATGATGAACCTGTCGCTACAAG 567
DB 340 GCTGCATCACTGACTACAGTATGCAATAGGTCATGATGATGAACTGGTTCGCGCAGCGT 281
QY 568 GCCCTGGTGGTATCGCA 584
DB 280 TCCATGATATCTGGCGTA 264

RESULT 3
BI879124/c      502 bp      mRNA      linear      EST 13-FEB-2002
LOCUS      BI879124
DEFINITION      f04e08.y1 Zebrafish adult retina cdna Danio rerio cDNA clone
IMAGE:4145367 5' similar to TR:Q9PMN4 Q9PMN4 RHODOPSIN.; mRNA
sequence.
ACCESSION      BI879124
VERSION      BI879124.1 GI:16086395
KEYWORDS      EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 502)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WASHU Zebrafish EST Project 1998
Unpublished
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
Library constructed by: Chandra Tucker and Gregory Niemi DNA
Sequencing by: Washington University Genome Sequencing Center Clome
distribution: RessourcenZentrumPrimarDatenbank, Berlin, Germany
(web address: www.rzpd.de)
Trace considered overall poor quality
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1..502
/organism="Danio rerio"
/mol_type="mRNA"
/strain="wild-type"
/db_xref="taxon:7955"
/clone="IMAGE:4145367"
/sex="mixed"
/dev_stage="1-2 years"
/lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"
/clone_lib="Zebrafish adult retina cdna"
/notes="Vector: Lambda ZAP II (phuescript SK-); Site_1:
EcORI; Site_2: SalI; This Zebrafish library was
constructed by Dr. Susan E. Brockerhoff (email:
sbrocker@u.washington.edu) RZPD library number: 760"
BASE COUNT      98 a 163 c 125 g 116 t
ORIGIN
Query Match      6.5%; Score 41.6; DB 12; Length 502;
Best Local Similarity 51.0%; Pred. No. 3.4;
Matches 98; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
```

QY 300 GGACATGATCATGCACACCCCGGTCGTCGCCCTGGTTCGGGAGAACAACTCTTCCCG 359

Db	484	GGCCATGAGCATGTGAAGCCACCACATGATGGCATGGTTTCGCCAGACCGAAAGTTGTTG	425
Qy	360	CTGCTGGGTAGCGCTCACCCCCACCGCTCGCAGCTAGGAACGCCAGCGTCCCCACACGAC	419
Db	424	CCAGTTTGAGTCCACCATCCAGGGCTCAATGCATGTACTCGAGGACCATAGCCCCAT	365
Qy	420	AATACAGCCGACGTGATTCGCCAGCTGTTTCACATCTCGCCTCGCGGATGAGACGGT	479
Db	364	CTCACCGCGCAGGGTGGCTAAGTAGCCTTACAGGTTGCAGTCTCAGTCGGCAGAACGAA	305
Qy	480	GCAGGACTGCAA	491
Db	304	GTAGCCGTGCAA	293

RESULT 4	AV835132	LOCUS	DEFINITION
AV835132	275 bp	mRNA	linear EST 22-JUN-2001
DEFINITION	AV835132 K. Sato unpublished cDNA library: Hordeum vulgare subsp. spontaneum top three leaves adult, heading stage Hordeum vulgare subsp. spontaneum cDNA clone bah24018, mRNA sequence.		
ACCESSION	AV835132		
VERSION	AV835132.1	GI:14527221	
KEYWORDS	EST.		
SOURCE	Hordeum vulgare subsp. spontaneum		
ORGANISM	Hordeum vulgare subsp. spontaneum		
REFERENCE	Sato, K.		
AUTHORS	Barley EST sequencing project in NIG and Okayama Univ		
TITLE	Unpublished		
JOURNAL	Contact: Kazuhiro Sato		
COMMENT	Research Institute for Bioresources Okayama University, Barley germplasm Center Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan Email: kazeato@rib.okayama-u.ac.jp, URL: http://www.rib.okayama-u.ac.jp/barley/ Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct submission; database: http://www.shigen.nig.ac.jp/barley/Barley.html.		

FEATURES	source
1..275	Location/Qualifiers
/organism="Hordeum vulgare subsp. spontaneum"	
/mol_type="mRNA"	
/cultivar="H602"	
/db_xref="taxon:77009"	
/clone="bah24018"	
/tissue_type="top three leaves"	
/dev_stage="adult, heading stage"	
/clone_lib="K. Sato unpublished cDNA library: Hordeum vulgare subsp. spontaneum top three leaves adult, heading stage"	
BASE COUNT	30 a 123 c 67 g 46 t 9 others
ORIGIN	

Query Match	6.4%;	Score 40.6;	DB 9;	Length 275;
Best Local Similarity	48.2%;	Pred. No. 5.1;		
Matches 109;	Conservative	0;	Mismatches 117;	Indels 0; Gaps 0;
Qy	306	GATCATGCACACCCCGGGTGGCTGGCCCTCGCTTCGGGAGAACAACTCTTCCCGCTGCTG	365	
Db	1	GGTCTTCGACGGAAACNCCGCCCTCTGCTCCGTTTGCTCTCTGCTCGGACCGCG	60	
Qy	366	GGTAGCGCTACACCCCAACGCTCGCAGCTAGGAACGCCAGCGTCCCCACACCAATACG	425	
Db	61	CGGCCAGCGGCACCTACCTNCTCTGCTCGTTTCGCCCTCCCTCTCGCGCCCCCTCG	120	
Qy	426	ACGCCAGGTGATTCGACGCTGTTTACCATCTCGCTTCGCCGGATGAGACGTTGACGA	485	
Db	121	ACGGCTGGCGTACGCTGGGTTGAGTTACATCTCTCGTCTCCGACGGGACGCTGGCGCA	180	

RESULT 6	551 bp	linear	EST 23-APR-2002
AV392165			
LOCUS	AV392165	<i>Chlamydomonas reinhardtii</i> C9	<i>Chlamydomonas reinhardtii</i>
DEFINITION			cDNA clone CM083e05 r 5' mRNA sequence.

	AV392165	Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii linear EST125-AY847200
LOCUS	AV392165	551 bp mRNA
DEFINITION	cDNA clone CM083e05_r_5', mRNA sequence.	
ACCESSION	AV392165	
VERSION	AV392165.1	GI:6546381
KEYWORDS	EST.	
SOURCE	Chlamydomonas reinhardtii	
ORGANISM	Chlamydomonas reinhardtii	
	Eukaryota; Viridiplantae;	
	Chlamydomonadaceae; Chlamydomonas.	
	Chlorophyta; Chlorophyceae; Volvocales;	

VERSION	AV392165.1	GI:6546381
KEYWORDS	EST.	
SOURCE	Chlamydomonas reinhardtii	
ORGANISM	Chlamydomonas reinhardtii	
	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;	
	Chlamydomonadaceae; Chlamydomonas.	

REFERENCE
1 (bases 1 to 551)
AUTHORS
Asanizumi, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE
A large scale structural analysis of cDNAs in a unicellular green
alga, *Chlamydomonas reinhardtii*. I. Generation of 3433
non-redundant expressed sequence tags
JOURNAL
DNA Res. 6 (6), 369-373 (1999)
MEDLINE
20152988
PUBMED
10691129

CONTACT: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
E-MAIL: yknaka@kazusa.or.jp
WWW: <http://www.kazusa.or.jp/en/plant/>

Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
E-mail: shimada@kazusa.or.jp, ip/en/plant/

108 CGAGCTCATCTCGTCAATTGTGCGGCGCACTGCCAATGAGGACGTCTGACGGACCT 167

QY 102 GCGCATGGCGTCCGGTCTCGGAGTCTGGAGCAGCGCGTGAACATATGCAACAGGGAATTTGCCGG 161

Db 168 GGCCTGGCGCGCGCGCGAGTGGAGGGCGGCTACGGCGCAGAGTCGTCAGCTGGCGC 227

CCGCTGGTGTCTGACCGTTCCAGCTTCGC 221

[illegible]

Db 228 CCGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCAGACCCAG 287

CCTGTCACGAACGACTGTCCA CTCAAG 281

QY
222 TATGAAGTGGCAACGTGTCCTGGATGACCTCATCGGTTCTC

DB
288 CTTGGCCGTCCGCTGCGTGGGCCACTCGCTGGCGGGGACCGCCGGCTGCCCTGTTCGAT 347

341

[illegible]

DB
348 CCTGATGCACCA CGACGAGGAGTTTGCGGCGGCATCTACGGCGGCGTGCCCATGCCCGG 407

368

[illegible]

Db 408 CAAGAAGCAAGGCAGCTACATGAT 434

RESULT. 7

RECEIVED
BI996341

LOCUS	BI996341	552 bp	linear	ESI 23-001
reinhardt11	1021027007	v2	C	Stress II (normalized),

DEFINITION 103103/A07.yz C: TERMINALIZE SS 103103, C:

```

QY 342 GGAGACAACTCTTCGGCTGCTGGGT 368
    |||||
Db 410 CAAGAAGACGAGGCGAGCTACATGAT 436

RESULT 8
B1727879
LOCUS B1727879 584 bp mRNA linear EST 19-SEP-2001
DEFINITION Lambda Zap II Chlamydomonas reinhardtii CC-1690, Stress II (normalized),
ACCESSION B1727879
VERSION B1727879.1 GI:15703574
KEYWORDS EST
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 584)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
JOURNAL Unpublished
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES
    source
    1..584
        /organism="Chlamydomonas reinhardtii"
        /mol_type="mRNA"
        /strain="CC-1690 wild type mt+ 21gr"
        /db_xref="taxon:3055"
        /clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
        ), Lambda Zap II"
        /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
        XhoI; Stress condition II library, constructed by John
        Davies and Jeffrey McDermott, combines cDNAs from CC-1690
        cells grown to mid-log phase in TAP (NH4+ - containing)
        and shifted to TAP - NO3- (24hrs); H2 production
        conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
        Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
        sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
        PolyA mRNA was purified from each sample, pooled and cDNA
        synthesized. The cDNA was directionally cloned into lambda
        Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
        sites. pBluescript II SK- plasmids were excised from the
        lambda Zap clones by superinfection with ExAesist
        (Stratagene) phage. The library was normalized using
        method 4 described in Bonaldo et al., (1996) Genome
        Research 6: 791-806."
BASE COUNT 106 a 188 c 197 g 93 t
ORIGIN

Query Match 6.4%; Score 40.6; DB 12; Length 584;
Best Local Similarity 45.3%; Pred. No. 6.3;
Matches 148; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 42 CGACCTCGTGGGTACATTCGGCTCGCGCGCCCTAGGGGGCGCTGCCAGGCGCCT 101
    |||||
Db 47 CGAGCTCATCTGTCATTTGCGCGGCACTGCCAACATGAGGAGCTGCTGACGACCT 106

QY 102 GCGCATGCGCTCGGGTTCTGAGGACGCGGTGAATGCAACAGGGAATTTGCCCGG 161
    |||||
Db 107 GCGCGCGCGCGCGAGTGGAGGCGGCTACGCGCAGAGTCTGCTGAGCTTGGCGC 166

QY 162 TTGCTCTTCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 221
    |||||
Db 167 CCGCAAGGTGTTTGACGAGATCAAGGAGTACGTGCTGAACCTCAAGGCCCCAGAACCCAG 226

QY 222 TTATGAAGTGGCAACGTTGTCGGGATGACCATGTACGAACGACTGCTCCAACCTCAAG 281
    |||||
Db 227 CTTGCGCGTCCGCTGCTGGGGCACTGCTGGCGGCGCACCGCGGCTGCTGCTGCTGAT 286

QY 282 CATTGCTGATGAGGAGGAGGACATCATGTCACACACCCCGGGTGGTGGCTGCTGCG 341
    |||||
Db 287 CTTGATGACCAACGAGGAGGAGTTCGCGCGGCGCATCTACGCGGCGGCTGCCATGCGCGG 346

QY 342 CGAGAACAACTCTTCGGCTGCTGGGT 368
    |||||
Db 347 CAAGAAGACGAGGCGAGCTACATGAT 373

RESULT 9
B1727879
LOCUS B1727879 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION clone CSODI015YB03 3-PRIME, mRNA sequence.
ACCESSION B1727879
VERSION B1727879.1 GI:30378083
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI015CA02NP1.

FEATURES
    source
    1..1201
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CSODI015YB03"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and cloned into the Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN

Query Match 6.4%; Score 40.4; DB 13; Length 1201;
Best Local Similarity 10.4%; Pred. No. 8.8;
Matches 52; Conservative 233; Mismatches 212; Indels 3; Gaps 1;

QY 34 GCTTCGCGGACTCGTGGGTACATTCGGCTGCGGCGCCCTAGGGGGCGCTGCC 93
    |||||
Db 618 GNTNTSSSSSTNNNNSSSSSSNNNTNTTBTBTSSSTSSSTSSSTSSSTSSST 677

QY 94 AGGCGCCTCGCGGCTGCGGTCTCGGAGGACGCGTCAACTATCAACAGGGAAT 153
    |||||
Db 678 SSSSSSBTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 737

QY 154 TTGCCCGGTGCTCTTCTATCTTCCTCTTGGCTTGTCTGCTGCTGCTGCTGCTGCT 213
    |||||
Db 738 TTSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 797

QY 214 GCTTCGCTTATGAGTGGCGAAGTGTCCGGATGTACCATGTACGAGCACTGCTCC 273
    |||||
Db 798 TBSMTSSSBTCTSSSSSSSBTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 954

```

Mon Dec 22 13:28:45 2003

```

QY 274 AACTCAGCATGTCTATGAGCGGACATGATCATGACACCCCGGGTGGTGGCC 333
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 855 TSSSTTTTSSSBSTSTSTSSBBSBTTTSTSTSTSTSTSTSTSTSTSTSSSS 914
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 334 TCGGTTCGGGAGAACAACTCTCCCGCTGCTGGGAGCGCTACCCCGCTCGAGCT 393
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 915 SSSSTTSSSBSTSTSTSSSSSTSTSSSSBBSBBSSTSSSSSSSSSSSSSS 974
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 394 AGGAACGCGAGCTCCACACGACGACATACACCCAGCTGATCCACGCTGTTCAC 453
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 975 STTBSSTTTTSSSBSTSTSTSSSTTTSSSTSSBBSBBSSTSSSSSSSTSS 1034
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 454 ATCTCGCTCGCGGATGAGCGGTGAGGACTGCAATGCTCAATCTATCCCGGCGAC 513
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1035 SNKKSSSSSSSSBBSBBSSTSSSSSSBBSBTTSTSTSSSTSTSTSSSSSSST 1094
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 514 ATAACGGGTACCGTATGCG 533
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1095 SBTSSSTTTTSSSATBSB 1114
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
BM918259 1162 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6611605 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485649
DEFINITION 5', mRNA sequence.
ACCESSION BM918259 GI:19368638
VERSION BM918259.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1162)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strauberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Distribution: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLOW2016 row: n column: 18
High quality sequence stop: 567.
Location/Qualifiers
1..1162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5485649"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_106"
/note="Organ: Blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."
BASE COUNT 224 a 499 c 240 g 198 t 1 others
ORIGIN
Query Match 6.3%; Score 40.2; DB 12; Length 1162;
Best Local Similarity 54.4%; Pred. No. 9.8;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 313 CACACCCCGGGTGGTGGCTCGGAGAACAACTCTCCGCTGGTGGTGGCG 372

```

```

Db 715 CCCCCCCCCGGTCTCTGCCACACCCCGGAGCCCAAAACCCCGGGGACCCCTCC 774
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 373 CTCACCCCGACGCTCGAGTAGGAGCGGACGCTCCCGACGACATACGAGCGCAC 432
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 775 CCCACTGCCACGCGACCCCGCCCATATCCGCCCGCCCTACCCGATCACCATTACCC 834
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 433 GTGATTCGCGAGCTGTTCACCATCTCGCC 461
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 835 GCCTGATCCCGGCTGTCACCCCGGCC 863
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
BX356664/c 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION Clone CSODI015B03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI015CA02NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI015B03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 116 a 88 c 93 g 398 t 506 others
ORIGIN
Query Match 6.3%; Score 40.2; DB 13; Length 1201;
Best Local Similarity 11.4%; Pred. No. 9.9; Indels 0; Gaps 0;
Matches 46; Conservative 196; Mismatches 171;

QY 227 AAGTGGCAACGCTGCTGGGATGTACCATGTCCAGAACGACTGCTCCAACTCAAGCTG 286
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1185 AASSVAASSSSSSSSSVSVAASAAATVAAAASAAASSSSSSVVSSSSAAASSSSS 1126
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 TGTATGAGCGAGCGGATGATCATGCACACCCCGGGTGGCTCGCTCGGAGGA 346
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1125 SAAAVVAASSVSATSSGAAAAAASAAVSAASSSSSSSSSSSSSSVSVAASVA 1066
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 ACAACTCTCCCGCTGCTGGTAGCGCTCACCCCGGCTCGCAGCTAGGACGCGAGG 406
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1065 AAATVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1006
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 TCCCGACGACAAATACGACGCGCATTCGATTCAGCTGTTTCCACATCTCCCTCGCC 466
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1005 AARASSVAAAVSSAAASVSSABAAAAAASVSSSSSSSSSSSSSSSSSSSSSS 946
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 GCGATGAGCGGTGCGAGCTGCAATTCGCTCAATCTATCCCGGCCACATACCGGTACC 526
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

Db      945 AVSASSSSSSAAVSSAAVSSSSVAAASSSSSSSSVAAASSSSSSAAVAVSAVAAA 886
QY      527 GTATGGCTGGGATATGATGAACTGGTGCCTTACAAAGCCCTGGTGGTATCGCAGC 586
Db      885 VSWSSVASSASVASVSSRRSSAANAASSAGSAAVAWAASSSSSSVAVVSSAAVSSMAS 826
QY      587 TGCTCCGATCGTGATCAGGCGCACACCATCACCACCATCA 629
Db      825 AAASSAAVSSSSSSSAGAVSSSAKSVASSAASVSSAGSSSA 783

RESULT 12
LOCUS   AL513886                1201 bp    mRNA    linear    EST 08-MAY-2003
DEFINITION AL513886 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA0062G08
5-PRIME, mRNA sequence.
ACCESSION AL513886
VERSION   AL513886.2 GI:30463771
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     1 (bases 1 to 1201)
JOURNAL   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
COMMENT   Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777380.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4924.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?req=CL0BA0062G08P1&cluster=4924.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BA0062G08P1.

FEATURES             source
    source
    1..1201
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CL0BA0062G08"
    /tissue_type="PLACENTA"
    /clone_lib="Homo sapiens PLACENTA"
    /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
    with a NotI-oligo(dT) primer. Five prime end enriched,
    double-strand cDNA was digested with Not I and cloned into
    the Not I and EcoRV sites of the pCMVSPORT 6 vector.
    Library was not normalized."
BASE COUNT    201 a 311 c 349 g 146 t 194 others
ORIGIN
    Query Match      6.3%; Score 40; DB 9; Length 1201;
    Best Local Similarity 26.5%; Pred. No. 11;
    Matches 103; Conservative 104; Mismatches 179; Indels 3; Gaps 1;

QY      32 GCGGCTTCGCCGACCTCGTGGGGTACATTCGCTCGCGCGCCCCCTAGGGGGCGGTG 91
Db      1129 GSGCSCGTGGCSBSNCTYKKGKBSSSCCSVSSSGSGSCSSCSGCGGGGGGGGG 1070
QY      92 CAGGCGCCCTGGCGCATGCGGTCTCGGGTCTTGAGAGACGCGGTGAACATGCAACAGGA 151
Db      1069 CCSCGKGGGGKSSSGSGGCCCGCGG---GGSSSSSGSGGCGSCWAAASSSAYKKGKK 1013
QY      152 ATTGCGCGGTGCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 211
Db      1012 GKXKTTTTWMAAATHTHTTWTWTWTITTTTCTTAAGGGGTAKVAKCCWCCGCCCAAGCTIS 953
QY      212 CAGCTTCCGCTTATGAAGTGCACACGCTGTCCGGGATGTACCATGTGCAGAACGACTGCT 271

```

```

Db      952 GACSCCCGCAADAACVCGAGMDSGAMKGTGVGSCCTTSRCKWGGGTTSGMMVGCATTY 893
QY      272 CCAACTCAAGCATTCGTGTATGAGGAGCGGAGCATCATGTCACACACCCCGCGGTGCGTGC 331
Db      892 AYBSYTGTRRTWTGTSBTCTYASGSGSMYSSKRKBGKCCMAYAACSCGAGASCST 833
QY      332 CCTGCGTTGGGAGAACAACTCTTCCCGTGTGGGTAGCGTCAACCCACGCTCGCAG 391
Db      832 SGCSKGTTKTTTKTGTCGTGAAGASMAABRTWAGGGGGGGGGCCCTCYSCMCCCCCYB 773
QY      392 CTAGGAAGCCGAGCGTCCCCACGACGA 420
Db      772 BBCCCMCHCTKCSCKMCCRGACTYCCCCA 744

RESULT 13
LOCUS   BJ252669/c                359 bp    mRNA    linear    EST 05-APR-2002
DEFINITION BJ252669 Y. Ogiwara unpublished cDNA library, Wh_f Triticum
aestivum cDNA clone whf25g19 3', mRNA sequence.
ACCESSION BJ252669
VERSION   BJ252669.1 GI:20061830
KEYWORDS EST.
SOURCE    Triticum aestivum (bread wheat)
ORGANISM  Triticum aestivum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
TITLE     1 (bases 1 to 359)
JOURNAL   Ogiwara,Y. and Murai,K.
COMMENT   Expressed genes in Triticum aestivum
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES             source
    source
    1..359
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="Chinese Spring"
    /db_xref="taxon:4565"
    /clone="whf25g19"
    /tissue_type="spike at flowering date"
    /dev_stage="freekes" scale 10.5.1"
    /clone_lib="Y. Ogiwara unpublished cDNA library, Wh_f"
BASE COUNT    70 a 115 c 107 g 67 t
ORIGIN
    Query Match      6.1%; Score 39; DB 12; Length 359;
    Best Local Similarity 58.0%; Pred. No. 14;
    Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY      24 CCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCCGCTCGCGCGCCCCCTAGS 83
Db      297 CTTCAAGTCAACAGCGCGCTCTGGAAGCGCTCAGGCGCGTGCAGCGCGTGCCTCGG 238
QY      84 GGGCGCTGCGGAGCGCCCTGGCGCATGCGCTCCGGGTTCTGGAGGAGCGGCTGAACCTATG 142
Db      237 GGACGCGCGACCCCTGGCGCAGACGCTGACAGTGTGCCGTTGCACGTGCCCAAGG 179

RESULT 14
LOCUS   BJ246716                375 bp    mRNA    linear    EST 05-APR-2002
DEFINITION BJ246716 Y. Ogiwara unpublished cDNA library, Wh_f Triticum
aestivum cDNA clone whf25g19 5', mRNA sequence.
ACCESSION BJ246716
VERSION   BJ246716.1 GI:20058228

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

REFERENCE
1 (bases 1 to 375)
Ogiwara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..375
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whf25g19"
/tissue_type="spike at flowering date"
/dev_stage="Peekes", scale 10.5.1"
/clone_lib="Y. Ogiwara unpublished cdna library, Wh_f"
BASE COUNT      81 a   107 c   110 g   77 t
ORIGIN
Query Match      6.1%; Score 39; DB 12; Length 375;
Best Local Similarity 58.0%; Pred. No. 14;
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 24 CCTTACATGGCGTTCCGACCTCGTGGGTACATTCCGCTCGTCGGGGCCCCCTAGG 83
Db 36 CTTCAAGTGAACAGCCCGCTCTGGAAGCGCTCAGGGCGGTGACGCGCGTGG 95

Qy 84 GGGCGCTGCCAGGCGCCCTGGCGCATGCGCTCGGGTTCTGGAGGAGCGCGTGAATG 142
Db 96 GGACGCGGAGCGCCCTGGGGCGGACGTCGACGTGCTGCGCGTGCACGTGCCCAAGG 154

RESULT 15
LOCUS      CG335916
DEFINITION OGUAJ60TV ZM.0.7.1.5 KB Zea mays genomic clone ZMMBma0393124,
genomic survey sequence.
ACCESSION  CC335916
VERSION    CC335916.1 GI:30805329
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick
, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek
, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"

FEATURES
source

```

```

/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0393124"
/clone_lib="ZM.0.7.1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
BASE COUNT      103 a   298 c   291 g   148 t
ORIGIN
Query Match      6.1%; Score 39; DB 29; Length 840;
Best Local Similarity 53.6%; Pred. No. 18;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 39 CGCGACCTCGTGGGTACATTCCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGC 98
Db 292 CCCGGCGTCGACGGGCCACTCCGCTTCACTCGCGCGCTCGCGGGAACGC 351

Qy 99 CCTGGCGCATGGGCTCCGGGTTCTGGAGGACGCGTGAATGCAACAGGGAATTGGC 158
Db 352 TGTGGCGCTGTCAGCTGCTCTGGGTGCGGGCGGACCCGGTAACCTCCGACTCCGC 411

Qy 159 CGGTGGCTCTTTCTCTATCTTCTCTTGGCT 189
Db 412 CGGTGGCTTGGCTGCGGACGTCATCTTTGTCT 442

Search completed: December 20, 2003, 06:54:51
Job time : 1638.15 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.8681 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-27
Perfect score: 636
Sequence: 1 ATGTTGGTAGGTCATCGA.....ATCACCACCATCACTAATAG 636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	3 US-08-612-973-27	Sequence 27, Appl
2	636	100.0	636	3 US-08-927-597-27	Sequence 27, Appl
3	597.4	93.9	606	3 US-08-612-973-25	Sequence 25, Appl
4	597.4	93.9	606	3 US-08-927-597-25	Sequence 25, Appl
5	597	93.9	723	3 US-08-612-973-21	Sequence 21, Appl
6	597	93.9	723	3 US-08-927-597-21	Sequence 21, Appl
7	556	87.4	561	3 US-08-612-973-23	Sequence 23, Appl
8	556	87.4	561	3 US-08-927-597-23	Sequence 23, Appl
9	515	81.0	795	3 US-08-612-973-5	Sequence 5, Appl
10	515	81.0	795	3 US-08-927-597-5	Sequence 5, Appl
11	512	80.5	2082	3 US-08-612-973-47	Sequence 47, Appl
12	512	80.5	2082	3 US-08-927-597-47	Sequence 47, Appl
13	512	80.5	2433	3 US-08-612-973-49	Sequence 49, Appl
14	512	80.5	2433	3 US-08-927-597-49	Sequence 49, Appl
15	456.4	71.8	636	3 US-08-612-973-7	Sequence 7, Appl
16	456.4	71.8	636	3 US-08-927-597-7	Sequence 7, Appl
17	451.6	71.0	636	3 US-08-612-973-13	Sequence 13, Appl
18	451.6	71.0	636	3 US-08-927-597-13	Sequence 13, Appl
19	448	70.4	1539	2 US-08-470-4268-17	Sequence 17, Appl
20	448	70.4	1863	2 US-08-470-4268-14	Sequence 14, Appl
21	443.2	69.7	9595	3 US-09-014-416-4	Sequence 4, Appl
22	443.2	69.7	9599	3 US-09-014-416-6	Sequence 6, Appl
23	441.6	69.4	932	1 US-08-081-072-15	Sequence 15, Appl
24	441.6	69.4	932	1 US-08-449-093A-15	Sequence 15, Appl
25	441.6	69.4	2116	3 US-08-191-160-21	Sequence 21, Appl
26	440	69.2	1037	1 US-08-462-195-1	Sequence 1, Appl
27	440	69.2	1037	2 US-08-636-883-1	Sequence 1, Appl

28	440	69.2	1037	3 US-09-127-829-1	Sequence 1, Appl
29	440	69.2	9472	4 US-08-150-204E-96	Sequence 96, Appl
30	428.8	67.4	1167	1 US-08-324-977-9	Sequence 9, Appl
31	428.8	67.4	1167	2 US-08-384-616-9	Sequence 9, Appl
32	428.8	67.4	1167	2 US-08-904-686A-9	Sequence 9, Appl
33	428.8	67.4	1167	3 US-09-315-850-9	Sequence 9, Appl
34	428.8	67.4	1499	1 US-08-324-977-3	Sequence 3, Appl
35	428.8	67.4	1499	2 US-08-384-616-3	Sequence 3, Appl
36	428.8	67.4	1499	2 US-08-904-686A-3	Sequence 3, Appl
37	428.8	67.4	1499	3 US-09-315-850-3	Sequence 3, Appl
38	428.8	67.4	6039	1 US-08-324-977-11	Sequence 11, Appl
39	428.8	67.4	6039	2 US-08-384-616-11	Sequence 11, Appl
40	428.8	67.4	6039	2 US-08-904-686A-11	Sequence 11, Appl
41	428.8	67.4	6039	3 US-09-315-850-11	Sequence 11, Appl
42	428.8	67.4	9030	1 US-08-324-977-13	Sequence 13, Appl
43	428.8	67.4	9030	2 US-08-384-616-13	Sequence 13, Appl
44	428.8	67.4	9030	2 US-08-904-686A-13	Sequence 13, Appl
45	428.8	67.4	9030	3 US-09-315-850-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-27
; Sequence 27, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLSCULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..633
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..630

		Query Match	100.0%;	Score 636;	DB 3;	Length 636;
		Best Local Similarity	100.0%;	Pred. No. 1.7e-157;	Mismatches 0;	Indels 0; Gaps 0;
		Matches 636;	Conservative 0;			
Qy	1	ATGTTGGGTAAAGTCATCGATAACCCCTTACATCGGCCTTCGCCGACTCGTGGGGTACATT	60			
Dd	1	ATGTTGGGTAAAGTCATCGATAACCCCTTACATCGGCCTTCGCCGACTCGTGGGGTACATT	60			
Qy	61	CCGCTCGTCGGCGCCCCCTTAGSGGGCGCTGCCAGGGCCCCTGGGGCATATGCGGT	120			
Dd	61	CCGCTCGTCGGCGCCCCCTTAGSGGGCGCTGCCAGGGCCCCTGGGGCATATGCGGT	120			
Qy	121	CTGGAGGACGGCGTGAACAATGCAACAGGGAAATTTGCCGGTTCCTTCCTATCTTC	180			
Dd	121	CTGGAGGACGGCGTGAACAATGCAACAGGGAAATTTGCCGGTTCCTTCCTATCTTC	180			
Qy	181	CTCTTGGCTTTTGCTGTCCTGTCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG	240			
Dd	181	CTCTTGGCTTTTGCTGTCCTGTCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG	240			
Qy	241	TCCGGATGTACCATGTCAACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGCACGG	300			
Dd	241	TCCGGATGTACCATGTCAACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGCACGG	300			
Qy	301	GACATGATCATGACACACCCCGGGTCGTCGCCCTCGGTCGGGAGAACAACTCTTCCCGC	360			
Dd	301	GACATGATCATGACACACCCCGGGTCGTCGCCCTCGGTCGGGAGAACAACTCTTCCCGC	360			
Qy	361	TGCTGGGTAGCGCTACACCCCGAGCTCGCAGGTAGGAACGCCAGCGTCCCCACCAGACA	420			
Dd	361	TGCTGGGTAGCGCTACACCCCGAGCTCGCAGGTAGGAACGCCAGCGTCCCCACCAGACA	420			
Qy	421	ATAACGACGCCAGCTCGATTCACAGCTGTTCACCATCTCGCTCCCGGCATGAGACGGTG	480			
Dd	421	ATAACGACGCCAGCTCGATTCACAGCTGTTCACCATCTCGCTCCCGGCATGAGACGGTG	480			
Qy	481	CAGGACTGCNAATTGCTCAATCTATCCGGGCACATAAAGGTCACCGTATGGCTTGGGAT	540			
Dd	481	CAGGACTGCNAATTGCTCAATCTATCCGGGCACATAAAGGTCACCGTATGGCTTGGGAT	540			
Qy	541	ATCATGATGAACCTGGTCGCTCAACAGGCCCTGGTGATATCGCAGCTGTCCGGATCGTG	600			
Dd	541	ATCATGATGAACCTGGTCGCTCAACAGGCCCTGGTGATATCGCAGCTGTCCGGATCGTG	600			
Qy	601	ATCGAGGCGACACCATCACCAACCATCATATAG	636			
Dd	601	ATCGAGGCGACACCATCACCAACCATCATATAG	636			

RESULT 2
US-08-927-597-27
; Sequence 27, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/08/927,597
5 FILING DATE:
6 CLASSIFICATION:
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US 08/612,973
9 FILING DATE: 11-MAR-1996
10 ATTORNEY/AGENT INFORMATION:
11 NAME: BYRNE, THOMAS E.
12 REGISTRATION NUMBER: 32,205
13 REFERENCE/DOCKET NUMBER: 1487-10
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: (703) 816-4000
16 TELEFAX: (703) 816-4100
17 INFORMATION FOR SEQ ID NO: 27:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 636 base pairs
20 TYPE: nucleic acid
21 STRANDEDNESS: single
22 TOPOLOGY: linear
23 MOLECULE TYPE: cDNA
24 HYPOTHETICAL: NO
25 ANTI-SENSE: NO
26 FEATURE:
27 NAME/KEY: CDS
28 LOCATION: 1..633
29 FEATURE:
30 NAME/KEY: mat_peptide
31 LOCATION: 1..630
32 US-08-927-597-27

```

Query Match	100.0%;	Score 636;	DB 3;	Length 636;
Best Local Similarity	100.0%;	Pred. No. 1.7e-157;	Indels 0;	Gaps 0;
Matches 636;	Conservative 0;	Mismatches 0;		
QY	1	ATGTTGGTAAAGTCATCGATACACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60	
DB	1	ATGTTGGTAAAGTCATCGATACACCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60	
QY	61	CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120	
DB	61	CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT	120	
QY	121	CTGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTTCTCTATCTTC	180	
DB	121	CTGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTTCCTTTCTCTATCTTC	180	
QY	181	CTCTTGGCTTTTGTGTCCTGCTCTGACCGCTTCAGCTTCCGCTTATGAAGTGGCAACGTG	240	
DB	181	CTCTTGGCTTTTGTGTCCTGCTCTGACCGTTTCAGCGTTTCAGCTTCCGCTTATGAAGTGGCAACGTG	240	
QY	241	TCCGGGATGTACATGTCAOGAAAGACTGCTCCAACTCAAGCATTTGCTATGAGCGACGG	300	
DB	241	TCCGGGATGTACATGTCAOGAAAGACTGCTCCAACTCAAGCATTTGCTATGAGCGACGG	300	
QY	301	GACATGATCATGCAACACCCCGGGTGGCTGCCCTCGTTCCGGAGAAACAACTCTTCCCGC	360	
DB	301	GACATGATCATGCAACACCCCGGGTGGCTGCCCTCGTTCCGGAGAAACAACTCTTCCCGC	360	
QY	361	TGCTGGGTAGCGCTCAACCCGCCACGCTCGCAGCTAGGAAACGCGGTCCTCCACACGACA	420	
DB	361	TGCTGGGTAGCGCTCAACCCGCCACGCTCGCAGCTAGGAAACGCGGTCCTCCACACGACA	420	
QY	421	ATACGACGCCAGTCGATTCACAGCTGTTTCAACAATCTCGCTCCGCGCATAGACGGTGG	480	
DB	421	ATACGACGCCAGTCGATTCACAGCTGTTTCAACAATCTCGCTCCGCGCATAGACGGTGG	480	
QY	481	CAGGACTGCAAATTTGCTCAATCTATCCGGGCCAATAACGGGTTCACCGTATGGCTGGGAT	540	
DB	481	CAGGACTGCAAATTTGCTCAATCTATCCGGGCCAATAACGGGTTCACCGTATGGCTGGGAT	540	

QY 541 ATGATGATGAAGTGGTCCCTACACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATCGTG 600
DB 541 ATGATGATGAAGTGGTCCCTACACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATCGTG 600
QY 601 ATCGAGGCGAGACACCATCACCACCATCACTAATAG 636
DB 601 ATCGAGGCGAGACACCATCACCACCATCACTAATAG 636

RESULT 3
US-08-612-973-25
; Sequence 25, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..603
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..600
; US-08-612-973-25

Query Match 93.9%; Score 597.4; DB 3; Length 606;
Best Local Similarity 99.8%; Pred. No. 2.1e-147;
Matches 598; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTTGGTAAAGTTCATGATACCCCTTACATCGCGCTTCGCGGACTGCTGGGGTACATT 60
DB 1 ATGTTGGTAAAGTTCATGATACCCCTTACATCGCGCTTCGCGGACTGCTGGGGTACATT 60
QY 61 CGCTCGTCGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120
DB 61 CGCTCGTCGGCGCCCTAGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT 120

QY 121 CTGGAGGACGGCGTGAAGTGAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGCGTGAAGTGAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTGCTGCTGACCCGTTCCAGCTTCCGGCTTATGAAGTGCACAACGTTG 240
DB 181 CTCTTGGCTTTGCTGCTGCTGCTGACCCGTTCCAGCTTCCGGCTTATGAAGTGCACAACGTTG 240
QY 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
DB 241 TCCGGGATGTACCATGTACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG 300
QY 301 GACATGATCATGCACACCCCGGGTGCCTCGCTTCCGGTTCGGGAGAACAACTCTTCCCGC 360
DB 301 GACATGATCATGCACACCCCGGGTGCCTCGCTTCCGGTTCGGGAGAACAACTCTTCCCGC 360
QY 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGGTCGCCACCAACGACA 420
DB 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGTAGGAACGCGGTCGCCACCAACGACA 420
QY 421 ATACAGCCCAAGTGCAGTTCAGCTGTTTCACTCTCGCTTCCGGCGGATGAGACGGTG 480
DB 421 ATACAGCCCAAGTGCAGTTCAGCTGTTTCACTCTCGCTTCCGGCGGATGAGACGGTG 480
QY 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATGCTTGGGAT 540
DB 481 CAGGACTGCAATTGCTCAATCTATCCCGGCCACATACCGGTCACCGTATGCTTGGGAT 540
QY 541 ATGATGATGAAGTGGTCCCTACACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATCGT 599
DB 541 ATGATGATGAAGTGGTCCCTACACAGCGCCCTGGTGGTATCGCAGCTGCTCCGGATCGT 599

RESULT 4
US-08-927-597-25
; Sequence 25, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:

Db 421 ATACGAGCCAGTCGATTCCAGCTGTTCCACATCTCCCTCGCGGCGATGAGACGGTG 480
Qy 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAATGCTGCGCTACACGGCCCTGGTGTATCCAGCTGCTCCGGATC 597
Db 541 ATGATGATGAATGCTGCGCTACACGGCCCTGGTGTATCCAGCTGCTCCGGATC 597

RESULT 6

US-08-927-597-21
; Sequence 21, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..717
; US-08-927-597-21

Query Match 93.9%; Score 597; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.7e-147;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTTGGGTAAGGTCAATCCATACCCCTTACATGCGGCTTCGCCAGCCCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGGTCAATCCATACCCCTTACATGCGGCTTCGCCAGCCCTCGTGGGGTACATT 60

Qy 61 CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCTTCGGGTT 120
Db 61 CCGCTCGTGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGCGCTTCGGGTT 120
Qy 121 CTGGAGGAGCGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
Db 121 CTGGAGGAGCGGCTGAACATATGCAACAGGGAATTTGCCGGTTGCTCTTTCTCTATCTTC 180
Qy 181 CTCTTGGCTTTGCTGCTGCTGTGACCGGTTCCAGCTTCCGCTTATGAAGTGCACACGTTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGTGACCGGTTCCAGCTTCCGCTTATGAAGTGCACACGTTG 240
Qy 241 TCCGGGATGTACCATGTACGAACGACCTCTCAACTCAAGCATTTGTGTATGAGGAGCG 300
Db 241 TCCGGGATGTACCATGTACGAACGACCTCTCAACTCAAGCATTTGTGTATGAGGAGCG 300
Qy 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCGCTTCCGCTTATGAAGTGCACACGTTG 360
Db 301 GACATGATCATGACACACCCCGGGTGGTGGCTTCCGCTTCCGCTTATGAAGTGCACACGTTG 360
Qy 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAACGCCAGGCTCCCAACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAACGCCAGGCTCCCAACACGACA 420
Qy 421 ATACGAGCCACGTCGATTTCCAGCTGTTTCCATCTCGCTCGCGGCGATGAGACGGTG 480
Db 421 ATACGAGCCACGTCGATTTCCAGCTGTTTCCATCTCGCTCGCGGCGATGAGACGGTG 480
Qy 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTGCAATGCTCAATCTATCCCGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGTATCGCAGCTGCTCCGGATC 597
Db 541 ATGATGATGAATGCTGCGCTACAAACGGCCCTGGTGTATCGCAGCTGCTCCGGATC 597

RESULT 7

US-08-612-973-23
; Sequence 23, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100

us-09-899-303a-27.rn1

Mon Dec 22 13:28:44 2003

```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555
; US-08-612-973-23

Query Match      87.4%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.4e-136;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGGCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGGCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGGCTCCGGGT 120
Db 61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGGCTCCGGGT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 CTCTTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGGATGTACCATGTCAACGACGATCTCAACTCAAGCATTTGTGTATGAGGACGG 300
Db 241 TCCGGGATGTACCATGTCAACGACGATCTCAACTCAAGCATTTGTGTATGAGGACGG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 360
QY 361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCCACACGACA 420
Db 361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCCACGTCGATTCACGATTCACCATCTCGCTCGCGCATGAGACGGTG 480
Db 421 ATACGACGCCACGTCGATTCACGATTCACCATCTCGCTCGCGCATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTTCGGGTTCGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTTCGGGTTCGGAT 540
QY 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556

RESULT 8
US-08-927-597-23
; Sequence 23, Application US/08927/597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..558
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..555
; US-08-927-597-23

Query Match      87.4%; Score 556; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.4e-136;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGGCGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGGTAAGGTGATCATGATACCCCTTACATGGGCTTCGGCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGGCTCCGGGT 120
Db 61 CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGGCTCCGGGT 120
QY 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
Db 121 CTGGAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTC 180
QY 181 CTCTTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 CTCTTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGGATGTACCATGTCAACGACGATCTCAACTCAAGCATTTGTGTATGAGGACGG 300
Db 241 TCCGGGATGTACCATGTCAACGACGATCTCAACTCAAGCATTTGTGTATGAGGACGG 300
QY 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 360
Db 301 GACATGATCATGCACACCCCGGGTGGTGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 360
QY 361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCCACACGACA 420
Db 361 TGCTGGGTAGGCTCACCCCCCAGCTCGAGCTAGGAAACGCCAGCGTCCCCACACGACA 420
QY 421 ATACGACGCCACGTCGATTCACGATTCACCATCTCGCTCGCGCATGAGACGGTG 480
Db 421 ATACGACGCCACGTCGATTCACGATTCACCATCTCGCTCGCGCATGAGACGGTG 480
QY 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTTCGGGTTCGGAT 540
Db 481 CAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTTCGGGTTCGGAT 540
QY 541 ATGATGATGAACCTGGT 556
Db 541 ATGATGATGAACCTGGT 556
```


Db 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTAGGAACGCGAGCGTCCCCACCAACGACA 420
Qy 421 ATACAGCCCACTGCTGATTTCCAGCTGTTACCATCTCGCCCTCGCGCGCATGAGACGGTG 480
Db 421 ATACAGCCCACTGCTGATTTCCAGCTGTTACCATCTCGCCCTCGCGCGCATGAGACGGTG 480
Qy 481 CAGGACTCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Db 481 CAGGACTCAATGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGGGAT 540
Qy 541 ATGATGATGAATGCTGCT 556
Db 541 ATGATGATGAATGCTGCT 556

RESULT 9
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789

US-08-612-973-5
Query Match 81.0%; Score 515; DB 3; Length 795;
Best Local Similarity 89.2%; Pred. No. 8.1e-126;
Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;
Qy 1 ATGTTGGGTAAGGTATCATGATACCCCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60

Db 1 ATGTTGGGTAAGGTATCATGATACCCCTTACATGCGGCTTCGCGGACCTCGTGGGTACATT 60
Qy 61 CCGCTCGTGGCGCGCCCTTAGGGGGCGCTCCAGAGGCGCTGGGGCGCATGGGGTTCGGGGTT 120
Db 61 CCGCTCGTGGCGCGCCCTTAGGGGGCGCTCCAGAGGCGCTGGGGCGCATGGGGTTCGGGGTT 120
Qy 121 CTGAGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTTCTCTATCTTC 180
Db 121 CTGAGAGGACGGCGTGAACATATGCAACAGGGAATTTGGCCGGTGTCTTTCTCTATCTTC 180
Qy 181 CTCTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGCG 240
Db 181 CTCTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGCG 240
Qy 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGCAGCG 300
Db 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCAATTTGTATGAGCAGCG 300
Qy 301 GACATGATCATGCACACCCCGGGTGTGCTGCTGCTGGGGAGAACAACTCTTCCCGC 360
Db 301 GACATGATCATGCACACCCCGGGTGTGCTGCTGCTGGGGAGAACAACTCTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTAGGAACGCGTCCCGCACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCCAACGCTCGCAGCTAGGAACGCGTCCCGCACACGACA 420
Qy 421 ATACGACGCCACGCTCGAT-----TCCAGCTGTTTCTGTTCCGCTATGTACGCG 480
Db 421 ATACGACGCCACGCTCGAT-----TCCAGCTGTTTCTGTTCCGCTATGTACGCG 480
Qy 439 -----TCCAGCTGTTTCTGTTCCGCTATGTACGCG 468
Db 481 GGGGACCTCTCGGATCTGTCTTCTGCTCTCCAGCTGTTTCCAGCTCTCGCGG 540
Qy 469 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGCCACATACGGGTACCGT 528
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGCCACATACGGGTACCGT 600
Qy 529 ATGCTTGGGATATGATGATGAACCTGCTCGCTCAACAGCGCCCTGCTGTTGATCGCAGCTG 588
Db 601 ATGCTTGGGATATGATGATGAACCTGCTCGCTCAACAGCGCCCTGCTGTTGATCGCAGCTG 660
Qy 589 CTCGGGATC 597
Db 661 CTCGGGATC 669

RESULT 10
US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

Mon Dec 22 13:28:44 2003

529 ATGGCTTGGATATCATGATGATGCTGCTACACGGCCCTGGTGGTATCGAGCTG 588
501 ATGGCTTGGATATCATGATGATGCTGCTACACGGCCCTGGTGGTATCGAGCTG 660
589 CTCGGGATC 597
661 CTCGGGATC 669

RESULT 11
US-08-612-973-47
; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-612-973-47

Query Match 81.0%; Score 515; DB 3; Length 795;
Best Local Similarity 89.2%; Pred. No. 8.1e-126;
Matches 597; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 1 ATGTTGGTAAAGTATCATGATGATGCTGCTACACGGCCCTGGTGGTATCGAGCTT 60
DB 1 ATGTTGGTAAAGTATCATGATGATGCTGCTACACGGCCCTGGTGGTATCGAGCTT 60
QY 61 CCGCTCTCGGGCGCCCTAGGGGGCGCTGCCAGGCGCTGGCGCATGCGTCCGGGT 120
DB 61 CCGCTCTCGGGCGCCCTAGGGGGCGCTGCCAGGCGCTGGCGCATGCGTCCGGGT 120
QY 121 CTGAGAGCGGTGAATATGCAACAGGGAATTTGCCGGTTCCTTTCTCTATCTTC 180
DB 121 CTGAGAGCGGTGAATATGCAACAGGGAATTTGCCGGTTCCTTTCTCTATCTTC 180
QY 181 CTCTTGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTCTTGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGGATGTACCATGTGTCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 TCCGGGATGTACCATGTGTCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 GACATGATCATGCAACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
DB 301 GACATGATCATGCAACCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
QY 361 TGCTGGTAGCGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 TGCTGGTAGCGCTCACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 ATACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 ATACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 439 -----TCCCGAGCTGTTTCCCGAGCTGTTTCCCGAGCTGTTTCCCGAGCTGTTT 468
DB 439 -----TCCCGAGCTGTTTCCCGAGCTGTTTCCCGAGCTGTTTCCCGAGCTGTTT 468
QY 481 GGGGACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GGGGACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 469 CATGAGACGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
DB 469 CATGAGACGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
QY 541 CATGAGACGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 CATGAGACGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Query Match 80.5%; Score 512; DB 3; Length 2082;
Best Local Similarity 89.2%; Pred. No. 6.4e-125;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGTAAAGTATCATGATGATGCTGCTACACGGCCCTGGTGGTATCGAGCTT 63
DB 4 TTGGTAAAGTATCATGATGATGCTGCTACACGGCCCTGGTGGTATCGAGCTT 63
QY 64 CTGCTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGGTTCGGGTTCTG 123
DB 64 CTGCTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGGCGCATGCGGTTCGGGTTCTG 123

Db 664 CGGATC 669

|||||

US-08-612-973-49

Sequence 49, Application US/08612973

Patent No. 6150134

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT

APPLICANT: BOSMAN, FONS

APPLICANT: DE MARTYNOFF, GUY

APPLICANT: BUYSE, MARIE-ANGE

TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHVE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/612,973

FILING DATE: 11-MAR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 2433 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2430

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 1..2427

US-08-612-973-49

Query Match 80.5%; Score 512; DB 3; Length 2433;

Best Local Similarity 89.2%; Pred. No. 6.6e-125; Indels 72; Gaps 1;

Matches 594; Conservative 0;

QY 4 TTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATCCG 63

Db 355 TTGGGTAAAGTTCATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGTACATCCG 414

QY 64 CTCGTGGCGCCCGCTAGCGGCGCTGCGAGGCGCTGGCGCATGGGTCGGGTTCTG 123

Db 415 CTCGTGGCGCCCGCTAGCGGCGCTGCGAGGCGCTGGCGCATGGGTCGGGTTCTG 474

QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTCTC 183

Db 475 GAGGACGGGTGAATATGCAACAGGGAATTTGCCCGGTTGCTCTTCTATCTCTC 534

QY 184 TTGGCTTTGCTGCTCTGCTGTCTGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGTGTC 243

Db 535 TTGGCTTTGCTGCTCTGCTGTCTGACCGCTTCAGCTTCGCTTATGAAGTGGCAACGTGTC 594

QY 244 GGGATGTACCATGTCCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 303

Db 595 GGGATGTACCATGTCCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGAC 654

QY 304 ATGATCATGCACACCGCGGCTGCGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC 363

Db 655 ATGATCATGCACACCGCGGCTGCGCTGCGTTCGGGAGAACAACTCTTCCCGCTGC 714

QY 364 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAAGCGCAGGTCGCCACACGACATA 423

Db 715 TGGGTAGCGCTCACCCCGACGCTCGCAGCTAGGAAAGCGCAGGTCGCCACACGACATA 774

QY 424 CGAGCGACGTCGAT-----TCCAGCTGTTTCCAGCTTCGCTCGCGGAT 471

Db 775 CGAGCGACGTCGATTTGCTGTTGGGCGGCTGCTTCTGTTCCGCTATGACGTGGG 834

QY 439 -----TCCAGCTGTTTCCAGCTTCGCTCGCGGAT 471

Db 835 GACCTCGCGGATCTGCTTCTTCCGCTCCAGCTGTTCCAGCTTCGCTCGCGGAT 894

QY 472 GAGACGCTGCGAGCTGCAATTGCTCAATCTATCCCGGACACATAACGGGTACCGTATG 531

Db 895 GAGACGCTGCGAGCTGCAATTGCTCAATCTATCCCGGACACATAACGGGTACCGTATG 954

QY 532 GCTTGGGATATGATGATGAACTGCTGCTTACAAACGCGCTTGGTGGTATCGCAGCTGCTC 591

Db 955 GCTTGGGATATGATGATGAACTGCTGCTTACAAACGCGCTTGGTGGTATCGCAGCTGCTC 1014

QY 592 CGGATC 597

Db 1015 CGGATC 1020

RESULT 14

US-08-927-597-49

Sequence 49, Application US/08927597

Patent No. 6245503

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT

APPLICANT: BOSMAN, FONS

APPLICANT: DE MARTYNOFF, GUY

APPLICANT: BUYSE, MARIE-ANGE

TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHVE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/927,597

FILING DATE: 11-MAR-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/612,973

FILING DATE: 11-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

```
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; NAME/KEY: mat_peptide
; LOCATION: 1..2427
US-08-927-597-49

Query Match      80.5%; Score 512; DB 3; Length 2433;
Best Local Similarity 89.2%; Pred. No. 6.6e-125;
Matches 594; Conservative 0; Mismatches 0; Indels 72; Gaps 1;

QY 4 TTGGTAAAGGTATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 63
DB 355 TTGGTAAAGGTATCGATACCCCTTACATCGCGCTTCGCCGACCTCGTGGGGTACATTCCG 414

QY 64 CTCGTGCGCGCCCTTAGGGGGCGCTGCGAGGCGCTGGCGCATGGGTCGCGGTTCTG 123
DB 415 CTCGTGCGCGCCCTTAGGGGGCGCTGCGAGGCGCTGGCGCATGGGTCGCGGTTCTG 474

QY 124 GAGGACGGGTGAATATGACACAGGGAATTTGCCGGTGTCTTCTATCTTCCTC 183
DB 475 GAGGACGGGTGAATATGACACAGGGAATTTGCCGGTGTCTTCTATCTTCCTC 534

QY 184 TTGGCTTTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTGCC 243
DB 535 TTGGCTTTGCTCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCAACGTGCC 594

QY 244 GGGATGTACCATGTTCACGAACGACTGCTCCAACTCAAGCAATGTGTATGAGCAGCGGAC 303
DB 595 GGGATGTACCATGTTCACGAACGACTGCTCCAACTCAAGCAATGTGTATGAGCAGCGGAC 654

QY 304 ATGATCATGCACACCCCGGGTGGTGGCTCGCTCGGTTCGGGAGAACAACTCTTCCCGTGC 363
DB 655 ATGATCATGCACACCCCGGGTGGTGGCTCGCTCGGTTCGGGAGAACAACTCTTCCCGTGC 714

QY 364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACAATA 423
DB 715 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAACGCCAGCGTCCCAACACGACAATA 774

QY 424 CGACGCCACGTGAT----- 438
DB 775 CGACGCCACGTGATTTGCTGTTGGGGCGGTCTTTCTGTTCCGCTATGATGTTGGG 834

QY 439 -----TCCAGCTGTTTCAACATCTGGCTCGCGGCAAT 471
DB 835 GACCTCTGCGGATCTGTCTTCTCGTCTCCAGCTGTTCCACCATCTCGCTCGCGGCAAT 894

QY 472 GAGAGCGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATG 531
DB 895 GAGAGCGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTACCGTATG 954

QY 532 GCTTGGGATATGATGATGAACCTGGTGCCTCAACAGCGCCCTGGTGTATCGCAGCTGCTC 591
DB 955 GCTTGGGATATGATGATGAACCTGGTGCCTCAACAGCGCCCTGGTGTATCGCAGCTGCTC 1014

QY 592 CGGATC 597
DB 1015 CGGATC 1020
```

RESULT 15
US-08-612-973-7

```
; Sequence 7, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNORF, GUY
; APPLICANT: BUYS, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..627
US-08-612-973-7

Query Match      71.8%; Score 456.4; DB 3; Length 633;
Best Local Similarity 86.8%; Pred. No. 1.7e-110;
Matches 545; Conservative 0; Mismatches 11; Indels 72; Gaps 1;

QY 1 ATGTTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT 60
DB 1 ATGTTGGGTAAAGTTCATCGATACCTTACATCGCGCTTCGCCGACCTCGTGGGTACATT 60

QY 61 CGGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTGCCGATGGCGTCCGGTT 120
DB 61 CGGCTCGTGGCGCCCCCTTAGGGGGGTGCTGCAGAGCCCTGGCGCATGGCGTCCGGTT 120

QY 121 CTGGAGGACGGGTGAACCTATGCAACAGGGAATTTGCCCGTTGCTCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAACCTATGCAACAGGGAATTTGCCCGTTGCTCTTCTCTATCTTC 180

QY 181 CTCCTTGGCTTCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGAACGTTG 240
DB 181 CTCCTTGGCTTCTGCTGCTGACCGTTCCAGCTTCGCTTATGAAGTGGCGAACGTTG 240

QY 241 TCCGGGATGTACCATGTGTACGAAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTGTACGAAACGACTGCTCCAACTCAAGCAATGTGTATGAGGACGG 300
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 19, 2003, 16:55:48 ; Search time 2385.36 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-29
Perfect score: 630
Sequence: 1 ATGGGTAGGTCATCGATAC.....TGATGATGAACCTGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rtd.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgtgo_hum.*
- 40: em_hgtgo_mus.*
- 41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	630	100.0	630	6	A48691	A48691 Sequence 29
2	630	100.0	630	6	AR157340	AR157340 Sequence
3	630	100.0	630	6	AX452778	AX452778 Sequence
4	630	100.0	630	6	AX685030	AX685030 Sequence
5	587.2	93.2	1546	14	HPCSTRUCTC	L12355 Hepatitis C
6	585.6	93.0	1786	14	HPCNZL1C	D14305 Hepatitis C
7	585.6	93.0	9390	14	HVCENS1	X76918 Hepatitis C
8	585.6	93.0	9456	14	HPCCEGS	D17763 Hepatitis C
9	582.4	92.4	1786	14	HPCUS114CE	D14309 Hepatitis C
10	576	91.4	1786	14	HPCTH85CE	D14307 Hepatitis C
11	576	91.4	9425	14	AF046866	AF046866 Hepatitis C
12	576	91.4	9454	14	HPCK3A	D28917 Hepatitis C
13	574.4	91.2	1786	14	HPCHEM26CE	D14311 Hepatitis C
14	539.4	85.6	541	6	A40621	A40621 Sequence 21
15	539.4	85.6	541	6	AX031599	AX031599 Sequence
16	539.4	85.6	541	6	AX031869	AX031869 Sequence
17	539.4	85.6	541	6	AX032139	AX032139 Sequence
18	539.4	85.6	541	6	BD172134	BD172134 New seque
19	539.4	85.6	541	14	HPCCOREED	D14599 Hepatitis C
20	533	84.6	541	6	A40619	A40619 Sequence 19
21	533	84.6	541	6	AX031597	AX031597 Sequence
22	533	84.6	541	6	AX031867	AX031867 Sequence
23	533	84.6	541	6	AX032137	AX032137 Sequence
24	533	84.6	541	6	BD172133	BD172133 New seque
25	504.2	80.0	541	6	A40613	A40613 Sequence 13
26	504.2	80.0	541	6	A40617	A40617 Sequence 17
27	504.2	80.0	541	6	AX031591	AX031591 Sequence
28	504.2	80.0	541	6	AX031595	AX031595 Sequence
29	504.2	80.0	541	6	AX031861	AX031861 Sequence
30	504.2	80.0	541	6	AX031865	AX031865 Sequence
31	504.2	80.0	541	6	AX032131	AX032131 Sequence
32	504.2	80.0	541	6	AX032135	AX032135 Sequence
33	504.2	80.0	541	6	BD172130	BD172130 New seque
34	504.2	80.0	541	6	BD172132	BD172132 New seque
35	504.2	80.0	541	14	HPCCOREEH	D14603 Hepatitis C
36	502.6	79.8	541	6	A40615	A40615 Sequence 15
37	502.6	79.8	541	6	AX031593	AX031593 Sequence
38	502.6	79.8	541	6	AX031863	AX031863 Sequence
39	502.6	79.8	541	6	AX032133	AX032133 Sequence
40	502.6	79.8	541	6	BD172131	BD172131 New seque
41	499.4	79.3	541	6	A40623	A40623 Sequence 23
42	499.4	79.3	541	6	A40625	A40625 Sequence 25
43	499.4	79.3	541	6	A40627	A40627 Sequence 27
44	499.4	79.3	541	6	AX031601	AX031601 Sequence
45	499.4	79.3	541	6	AX031603	AX031603 Sequence

ALIGNMENTS

RESULT 1

A48691
LOCUS A48691
DEFINITION Sequence 29 from Patent WO9604385.
ACCESSION A48691
VERSION A48691.1 GI:2302404
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 630)
AUTHORS Maertens, G., Bosman, F., De M.G. and Buysse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 29 15-FEB-1996;

linear PAT 07-MAR-1997

Mon Dec 22 13:28:45 2003

LOCUS AR157340 Sequence 29 from patent US 6245503. 630 bp DNA linear PAT 17-OCT-2001

DEFINITION AR157340 Sequence 29 from patent US 6245503.

ACCESSION AR157340 GI:16218274

VERSION AR157340.1

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 630)

AUTHORS Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.

TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use

JOURNAL Patent: US 6245503-A 29 12-JUN-2001;

FEATURES Location/Qualifiers

source 1..630

BASE COUNT 127 a 175 c 168 g 160 t

ORIGIN

Query Match 100.0%; Score 630; DB 6; Length 630;

Best Local Similarity 100.0%; Pred. No. 4.5e-165;

Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGTAAGGTCATCGATACCCCTTACGTGGGATTCGCGGATCTCATGGGTACATCCCG 60

Db 1 ATGGGTAAGGTCATCGATACCCCTTACGTGGGATTCGCGGATCTCATGGGTACATCCCG 60

Qy 61 CTGCTCGGCGCTCCGTPAGGAGCGCTCGCAAGAGCCCTTTCGCGATCGCGTGGGCGCTT 120

Db 61 CTGCTCGGCGCTCCGTPAGGAGCGCTCGCAAGAGCCCTTTCGCGATCGCGTGGGCGCTT 120

Qy 121 GAAGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTTTCTATTTTCCTT 180

Db 121 GAAGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTTTCTATTTTCCTT 180

Qy 181 CTGCTCTGTTCTCTTCTGCTTAATTCATCCAGCAGTCTAGTCTAGAGTGGCGGAATACGTCT 240

Db 181 CTGCTCTGTTCTCTTCTGCTTAATTCATCCAGCAGTCTAGTCTAGAGTGGCGGAATACGTCT 240

Qy 241 GGCCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTTGTACGAGGCGGATGAC 300

Db 241 GGCCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTTGTACGAGGCGGATGAC 300

Qy 301 GTTATCTGACACACCCGGCTGCATACCTTGTCCAGGACGCAATATACATCCAGTGC 360

Db 301 GTTATCTGACACACCCGGCTGCATACCTTGTCCAGGACGCAATATACATCCAGTGC 360

Qy 361 TGGACCCAGTGACACCTTACAGTGGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420

Db 361 TGGACCCAGTGACACCTTACAGTGGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420

Qy 421 CGCAGTCATGTGGACCTTATAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

Db 421 CGCAGTCATGTGGACCTTATAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

Qy 481 GACATGTGGGCGTGTCTTCTCTGCGGACAAAGCCTTCAAGTCAAGTCAAGTCAAGTCAAGT 540

Db 481 GACATGTGGGCGTGTCTTCTCTGCGGACAAAGCCTTCAAGTCAAGTCAAGTCAAGTCAAGT 540

Qy 541 CAAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCGCATCTTTCAGGACATCGAATG 600

Db 541 CAAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCGCATCTTTCAGGACATCGAATG 600

Qy 601 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630

Db 601 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630

RESULT 3

AX452778 AX452778 Sequence 29 from Patent EP1211315.

LOCUS AX452778 630 bp DNA linear PAT 06-JUL-2002

DEFINITION AX452778

ACCESSION AX452778

INNOGENETICS NV (BE)

Other publication CA 2172273 960215

Other publication AU 3382495 960304.

Location/Qualifiers

1..630

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

1..627

/note="unnamed protein product"

/codon_start=1

/protein_id="CAA03138.1"

/db_xref="GI:2302405"

/translation="MGKVIDTLTCGFADLMGVPLVAGPVGVARALAHGVRLEDDGI

NFATNLPCSFISIFLLALPSCLHPASLEWNTSGLYLITNDCSNSIVYEADVDI

LHTPGICPQDGNSTWTVPVAVKYVGATTAIRSHVLLVGAATWCSALYVG

DMCGAVFLVGOAFTFRPRRHQTVQCNCSLYPHLSGHRMAMDMMNW"

mat_peptide 1..624

/products="unnamed"

BASE COUNT 127 a 175 c 168 g 160 t

ORIGIN

Query Match 100.0%; Score 630; DB 6; Length 630;

Best Local Similarity 100.0%; Pred. No. 4.5e-165;

Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGTAAGGTCATCGATACCCCTTACGTGGGATTCGCGGATCTCATGGGTACATCCCG 60

Db 1 ATGGGTAAGGTCATCGATACCCCTTACGTGGGATTCGCGGATCTCATGGGTACATCCCG 60

Qy 61 CTGCTCGGCGCTCCGTPAGGAGCGCTCGCAAGAGCCCTTTCGCGATCGCGTGGGCGCTT 120

Db 61 CTGCTCGGCGCTCCGTPAGGAGCGCTCGCAAGAGCCCTTTCGCGATCGCGTGGGCGCTT 120

Qy 121 GAAGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTTTCTATTTTCCTT 180

Db 121 GAAGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTTTCTATTTTCCTT 180

Qy 181 CTGCTCTGTTCTCTTCTGCTTAATTCATCCAGCAGTCTAGTCTAGAGTGGCGGAATACGTCT 240

Db 181 CTGCTCTGTTCTCTTCTGCTTAATTCATCCAGCAGTCTAGTCTAGAGTGGCGGAATACGTCT 240

Qy 241 GGCCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTTGTACGAGGCGGATGAC 300

Db 241 GGCCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTTGTACGAGGCGGATGAC 300

Qy 301 GTTATCTGACACACCCGGCTGCATACCTTGTCCAGGACGCAATATACATCCAGTGC 360

Db 301 GTTATCTGACACACCCGGCTGCATACCTTGTCCAGGACGCAATATACATCCAGTGC 360

Qy 361 TGGACCCAGTGACACCTTACAGTGGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420

Db 361 TGGACCCAGTGACACCTTACAGTGGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420

Qy 421 CGCAGTCATGTGGACCTTATAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

Db 421 CGCAGTCATGTGGACCTTATAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

Qy 481 GACATGTGGGCGTGTCTTCTCTGCGGACAAAGCCTTCAAGTCAAGTCAAGTCAAGTCAAGT 540

Db 481 GACATGTGGGCGTGTCTTCTCTGCGGACAAAGCCTTCAAGTCAAGTCAAGTCAAGTCAAGT 540

Qy 541 CAAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCGCATCTTTCAGGACATCGAATG 600

Db 541 CAAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCGCATCTTTCAGGACATCGAATG 600

Qy 601 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630

Db 601 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630

RESULT 2

AR157340


```

VERSION      AX452778.1  GI:21712463
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 1211315-A 29 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              source
                1..630
                /organism="Hepatitis C virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:11103"
                1..627
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD38076.1"
                /db_xref="GI:21712464"
                /translation="MGKVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRALBDGI
                NFAFNLPCCSFSIFLLALFSLIHPAASLEWRNTSGLYVLNDCSNSIYVEADVI
                LHTPGCIPCVODGNTSTCWTPTVPTVAKYVGATTASIRSHVDLLVGAATMCSALYVG
                DMCAGVFLVQAFTFRPRRHQVQTCNCSLYPHLSGHRMADMMNNW"
              mat_peptide
                1..624
                /product="unnamed"
BASE COUNT   127 a 175 c 168 g 160 t
ORIGIN
Query Match      100.0%; Score 630; DB 6; Length 630;
Best Local Similarity 100.0%; Pred. No. 4.5e-165;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTAAAGTCATCGATACCCCTTACGTCGGGATTCGCGATCTCATGGGGTACATCCCG 60
DB 1 ATGGGTAAAGTCATCGATACCCCTTACGTCGGGATTCGCGATCTCATGGGGTACATCCCG 60
QY 61 CTCGTGGCGCTCCCGTAGAGCGCTGCGAAGAGCCCTTGCATGGCGTGGAGGCCCTT 120
DB 61 CTCGTGGCGCTCCCGTAGAGCGCTGCGAAGAGCCCTTGCATGGCGTGGAGGCCCTT 120
QY 121 GAAGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCCTT 180
DB 121 GAAGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCCTT 180
QY 181 CTCGCTCTGTTCTCTTGTCTTAATTCATCCAGCAGTAGCTAGAGTGGCGGAATACGCT 240
DB 181 CTCGCTCTGTTCTCTTGTCTTAATTCATCCAGCAGTAGCTAGAGTGGCGGAATACGCT 240
QY 241 GGCCTCTATGTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
DB 241 GGCCTCTATGTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAC 300
QY 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGC 360
DB 301 GTTATTCTGCACACCCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGC 360
QY 361 TGGACCCCAAGTGCACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACACCCGCTTCGATA 420
DB 361 TGGACCCCAAGTGCACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACACCCGCTTCGATA 420

```

421	QY	CGCAGTCATGTGGACCTATTAGTGGGCGCGCCACGATGTGCTCTGGCTCTACGTGGCT	480
421	Db	CGCAGTCATGTGGACCTATTAGTGGGCGCGCCACGATGTGCTCTGGCTCTACGTGGCT	480
481	QY	GACATGTGTGGGCTGTCTTCTCGTGGGCAAGCCCTTCACGTTTCAGACCTCGTCGCCAT	540
481	Db	GACATGTGTGGGCTGTCTTCTCGTGGGCAAGCCCTTCAGTTTCAGACCTCGTCGCCAT	540
541	QY	CAAAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG	600
541	Db	CAAAACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATG	600
601	QY	GCTTGGGATATCATGATGAACCTGTTAATAG	630
601	Db	GCTTGGGATATCATGATGAACCTGTTAATAG	630
RESULT 5			
LOCUS	HPCTRUCTC	1546 bp ss-RNA	linear VRL 27-JUL-1994
DEFINITION	Hepatitis C virus core and envelope proteins gene, 5' end of cds.		
ACCESSION	U12355		
VERSION	L12355.1	GI:410169	
KEYWORDS	core protein; envelope protein; structural region.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1 (sites)		
AUTHORS	Han, J.H., Shyamala, V., Richman, K.H., Brauer, M.J., Irvine, B., U-dea, M.S., Tekamp-Olson, P., Kuo, G., Choo, Q.L. and Houghton, M.		
TITLE	Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5' untranslated region and poly(A) tails at the 3' end		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	88 (5), 1711-1715	(1991)
MEDLINE	91156678		
PUBMED	1705704		
REFERENCE	2 (sites)		
AUTHORS	Choo, Q.-L., Richman, K., Han, J.H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P.J., Weiner, A., Bradley, D.W., Kuo, G. and Houghton, M.		
TITLE	Genetic organization and diversity of the hepatitis C virus		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	88 (6), 2451-2455	(1991)
MEDLINE	91172826		
PUBMED	1848704		
REFERENCE	3 (bases 1 to 1546)		
AUTHORS	Li, J.S., Vitvitski, L., Tong, S.P. and Trepo, C.		
TITLE	Identification of the third major genotype of hepatitis C virus in France		
JOURNAL	Biochem. Biophys. Res. Commun.	199 (3), 1474-1481	(1994)
MEDLINE	94197744		
PUBMED	8147893		
COMMENT	Original		
FEATURES	source text: Hepatitis C virus cDNA to genomic RNA.		
	Location/Qualifiers		
source	1..1546		
	/organism="Hepatitis C virus"		
	/db_xref="taxon:11103"		
	/tissue_type="serum"		
	/note="specific host (French) "		
	314..>1546		
	/standard_name="structural region"		
	/note="encodes core and envelope proteins"		
	/codon_start=1		
	/protein_id="AA020155.1"		
	/db_xref="GI:410170"		
	/translation="MSTLTPKQRTKNTIRPDVKKFPQGVQGVVYVLRGRPL		
	GVRTKRTSEMSQPRGRQPIKARSDGRSSWAGPWPYLYGCGCGWAGWLLSPRG		
	SRPSSGPNPDRRRNSKNGKIDITLTCTGADLMGYPLVAGVGVGVARALAHGVRALED		
	GINTFATGLPCVCSIFULLAFPSCLHPHFAASLEWRNTSGLYILTDCSSSIYVEAD		
	VILHTPGVCVQDGTSTCTPVTPTVPAVRVGTASTIRSHVDLLVGNATWCSALY		
	VQDMCGAVFLVQGAFFPRPHQTVCNCSLYFGHLSGRHAWDMWMSPAVGMV		
	AVHLRLPOTFLDIIAGAHGILAGLAYSVMQGNWAKVLIIMWPSGVDAGTRITGGSA		
CDS			

REFERENCE 2 (bases 1 to 1786)
AUTHORS Okamoto, H.
JOURNAL Unpublished
COMMENT Submitted (28-JAN-1993) to DDBJ by:
Hiroaki Okamoto
Immunology Division
Jichi Medical School
Kawachi
Tochigi 329-04
Japan
Phone: 0285-44-2111x3334
Fax: 0285-44-1557
Location/Qualifiers
1..1786
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="NZL1"
/db_xref="taxon:11103"
-1..277
/citation=[1]
278..>1786
/codon_start=1
/product="polyprotein precursor"
/protein_id="BAA03251.1"
/db_xref="GI:471114"
/translation="MTSLPKPQRYKNTIRPQDKFPGGQIVGGVTVLPRRPRRL
GVARTKTSERSQPRRQPIPKARRSERSWAQFYPWPLYNGECGWAGLLSPRG
SRPSWGNDRPRRSRLKGVLDLTCGFADLMGYIPLVGAPVGGVARALAHGVRALED
GINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSGLYVLTNDCSNSIYVBAEDD
VILHTPGVCPCVDGNTSTCTWPTVAVRVGATTASIRSHVDLLVGAATMCSALY
VGDMGCAFLVQCATFFPRRHQTQVTCNCSLYPGHLSHRMANMNMWSPAVGNV
AHVLRPQLFLDIIAGAHWGLIAGLYSMQGNWAKVAVIMVWFSGVDAHTYTTGGTA
SRHTQAFGLFDIGPQKQLQVNTGNSWHINSTALNCNCSINTGFIAGLYFYHKFNST
GCPQLSSCKPITFFRQWGLTDANITGPSDDRYCWHYAPRPCDIIVPASSV"
278..850
/product="C"
851..1426
/product="E"
1427..>1786
/product="E2/NS1"
BASE COUNT 366 a 521 c 497 g 402 t
ORIGIN

Query Match 93.0%; Score 585.6; DB 14; Length 1786;
Best Local Similarity 96.2%; Pred. No. 1.2e-152;
Matches 600; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 2 TCGGTAAGTCTACGATACCTTACGTGGGATTCCCGATCTCATGGGGTACATCCCGC 61
DB 633 TGGGTAAGTCTACGATACCTTACGTGGGATTCCCGATCTCATGGGGTACATCCCGC 692
QY 62 TCGTGGGCTCCGTAGGAGCGTCGCAAGAGCCCTTCGGATGCGGTAGGCGCCCTTG 121
DB 693 TCGTGGGCTCCTGTAGGAGCGTCGCAAGAGCCCTTCGGATGCGGTAGGCGCCCTTG 752
QY 122 AAGACGGGATAAATTTGCAACAGGAAATTTGCCGGTGTCTCTTTCTATTTCCTTC 181
DB 753 AAGACGGGATAAATTTGCAACAGGAAATTTGCCGGTGTCTCTTTCTATTTCCTTC 812
QY 182 TCGCTCTGTTCTTGTCTTAATTCACGAGCTAGTCTAGAGTGGCGGAATACGTCGTG 241
DB 813 TTGCTCTGTTCTTGTCTTAATTCACGAGCTAGTCTAGAGTGGCGGAATACGTCGTG 872
QY 242 GCCTCTATGCTCTTACCAAGCTGTTCCTAATAGCAGTATTTGTACGAGCGCGATGAGC 301
DB 873 GCCTCTATGCTCTTACCAAGCTGTTCCTAATAGCAGTATTTGTATGAGCGCGATGAGC 932
QY 302 TTATTCGACACACCGCGCTGCATACCTTGTGTCCAGACGCGCAATACATCCAGTGTCT 361
DB 933 TCATTCGACACACCGCGCTGTGTACCTTGTGTCCAGACGCGCAATACATCTACGTGTCT 992
QY 362 GGACCCCATGACACCTTACGAGTGTCAAGTACGTCGAGCAACACCGCTTTCGATAC 421

Db 993 GGACCCCATGACACCTTACGAGTGTGGAGTACGTCGGAGCAAACTACTGCTTCGATAC 1052
QY 422 GCAGTCATGTGACCTATTAGTGGCGGGCCACGATGTGCTCTCGCTCTACGTGGGTG 481
Db 1053 GCAGTCATGTGACCTATTAGTGGCGGGCCACGATGTGCTCTCGCTCTACGTGGGTG 1112
QY 482 ACATGTGTGGGCTGTCTCTCTCGTGGGACACGCTTACGTTTACAGCTCTCGCCATC 541
Db 1113 ATATGTGTGGGCTGTCTCTCTCGTGGGACACGCTTACGTTTACAGCTCTCGCCATC 1172
QY 542 AAACGTCGACACCTTAACTCTCGTGTACCCAGGCCATCTTTTACGAGCATCGAATCG 601
Db 1173 AAACGTCGACACCTTAACTCTCGTGTACCCAGGCCATCTTTTACGAGCATCGAATCG 1232
QY 602 CTTGGGATATGATGAACTGGT 625
Db 1233 CTTGGGATATGATGAACTGGT 1256

RESULT 7
LOCUS HVCENS1 9390 bp RNA linear VRL 17-JAN-1995
DEFINITION Hepatitis C virus genes for core, envelope and NS1 proteins.
ACCESSION X76918
VERSION X76918.1 GI:633201
KEYWORDS core protein; envelope protein; NS1 protein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; asRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9390)
AUTHORS Seelig, R., Weber, P., Seeling, H.P., Ledger, N., Bottner, C. and
Renz, M.
TITLE Hepatitis C virus type V genome isolated from a patient in Germany
JOURNAL Unpublished
AUTHORS Seelig, R.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1993) R. Seelig, Inst. of Immunol. and Mol.
Genetics, KriegsstraÙe 99, D- 76133 Karlsruhe, FRG
REFERENCE 3 (bases 1 to 9390)
AUTHORS Seelig, R.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-1995) R. Seelig, Inst. of Immunol. and Mol.
Genetics, KriegsstraÙe 99, D- 76133 Karlsruhe, FRG
COMMENT On Jan 24, 1995 this sequence version replaced gi:506489.
FEATURES
source
1..9390
/organism="Hepatitis C virus"
/mol_type="unassigned RNA"
/db_xref="taxon:11103"
1..2267
/organism="Hepatitis C virus"
/mol_type="unassigned RNA"
/sub_species="type V-D (Germany)"
/db_xref="taxon:11103"
/map="297 to 1970 nucleotide number"
/tissue_type="serum"
299..9364
/note="viral coding region"
/codon_start=1
/protein_id="CAA54244.1"
/db_xref="GI:633202"
/db_xref="SPTREMBL:Q68870"
/translation="MTSLPKPQRYKNTIRPQDKFPGGQIVGGVTVLPRRPRRL
GVARTKTSERSQPRRQPIPKARRSERSWAQFYPWPLYNGECGWAGLLSPRG
SRPSWGNDRPRRSRLKGVLDLTCGFADLMGYIPLVGAPVGGVARALAHGVRALED
GINFATGNLPGCSFSIFLLALFSLIHPAASLEWNTSGLYVLTNDCSNSIYVBAEDD
VILHTPGVCPCVDGNTSTCTWPTVAVRVGATTASIRSHVDLLVGAATMCSALY
VGDMGCAFLVQCATFFPRRHQTQVTCNCSLYPGHLSHRMANMNMWSPAVGNV
SHVLRPQLFLDIIAGAHWGLIAGLYSMQGNWAKVAVIMVWFSGVDAHTYTTGSA
AHGVTSLTSLFSSGPQKQLQVNTGNSWHINSTALNCNCSINTGFIAGLYFYHKFNST
GCPQLSSCKPITFFRQWGLTDANITGPSDDRYCWHYAPRPCDIIVPASSV"

CFTSPVVVGGTTDRKGVPTNWNENSDVFLLESRLPSPGRWFCGAMNNTGFLKTCG
APPNNYGGGNNPNNHSLFCTPTDCFRKATDYSRCGAGFWLTPRCMDVDPYRLWHY
PCTVNFITLTKVRFVGGFEHPTFAACNWTGRCERNIEBDRSEOHPLHSTTEALILP
CSFTMPALSTGLIHLHONIYDVQYLKVGSGNMGWALKWFEVLIIELLIADARCVVA
LWMLISQAEALLENLTKAAAGTHGILGVLVAFCAAHVGRKLVPLVITYSLTG
LWSLALLVLLPQRAYMSGDSATLGRVLVLFPGFTLSPWYHGLRILWNQYTI
CRCSALQVWPVPIIARGSDVILITSLIYPSLIFDITKLLIAGVLGSLYLIQASITA
TPYFVRAVLRLCMLVRSMMGGKYFOMIILSVGRWFNTVLYDLHAPQHWAAAGWR
LAVATEVLPSPMETKVTMGADTAACDILGLVSGARLGRVLLVADDDYRBMGR
LIAPITAQOTRGLLGIITVGLTGRDNVTVGEVQLSTATQCTFLGTVGVVMVIV
HGASRTLAGAKPALQITNVQDLVGNWAPPAGKSLPCACGSADLYLVTRDADVI
PARRGDSTASLLSPRLACILKSGGGVCMPCPSHAGIFRAAVCTRGVAKALQFIV
ETLSTQARSFMSNDSTPPAVPTQVGYLHAPTSGSKSTKPAAYVAQGNVILNP
SVAATLHGFSFMSRAYGTDPTNRTGNRTVTTGAKLTYSTYKGLFLADGCGSGADYLI
CDECAQDAPATILGIGTLDQAEAGVRLTVLATATPGSITVPHRSNIEEVALSGE
IFPYGKAATIALLGKRLHILFCHSKKDEIASLKLGMNLNAVAYHYRILVSPVTTG
DVVCAATDALMTGTFDSDVDNCVAVQYVDSLPDTPFSETRTAPODAVRSORR
GTRGRGLTYRYVAPGRPSGDFSVLCEBIDAGCSWYDLQPAETTVRLRAVLSIP
GLPVCQDHLDFWESVPTGLTHIDAHFLSQTKQOGLNFSILTAYQATVCARQAQPPSW
DEWKLVRKLTPLTLYRLGVPQNETCLTHPVTKYIMACMSADELVTTSTVWL
LGVLAAALAYCLSVGVVIVGHIELGKPALVPDKEYLYQYDEMEESQAPYIEQ
AQALHOFKVLGLQATQOAVIEPIVATNMOKLEAFWHKHMNFVSGIYQIAGL
PILGNPAVASLMAPTAVSTSPLTNOTMFEINLGSELPTTEDLVNLLPAILSPGALV
AAIGGLGRLDILLIAGGAGVSGALVAFKIMGELPTTEDLVNLLPAILSPGALV
GVICAILLRHHVGGEGAVQMMNRLIFAFSGNHVSPLHYTHYVPSADAAKVTALLSLLT
VTSLLRLHQWINEIDYSPCDDWLRITIMWVCSVLADFKTWLSAKIMFALPLGLFIS
CQKYGVWNGDGVITTRCPGATITGHVKNMSRLAGRTCANWYGTFFINEITIG
PSPPCSPNVTALWVAANSYVEVRVGDHYITGATEDELKCPQVPAAEFFTEVD
GVRLHYAPCKPLRDDITFMVGLNSYAIQSOLPCEPEPDVSVLTMRLDPFSHTAE
TAARLIARSGPSEASSASQLSA PSLKATCOTHPDPAELVDANLLWRQBMGSNIT
RVESRTILDSFEPLRAQTDIAELSVAAECFKPKPKYPALPITWAPDNPPLDLR
WKTDPDVPVTHCALPPRAGVPPPPRRKRTIQLDGSNVSAALAAKSPSVNPO
DENSSSGVDVDTQSTTSKVPVPGSDSESCSPMPEGEPPDPLDCDSWSTVSDS
EBOGVVCSMSYMTGALITPCSAEKEKLPISPLNSLLRHLNLYSTSSRASORQK
KVTDFRLOVLDHYKTLKEVRAKSRVARKMLTIEACALVPHSARSKFGYSADKV
RSLSSRAIDORSVSWDLEDDTTPPTIMAKNEVFCVDPARGKRRARLIVYDFRG
RVVCEKRVLDVILQKLSIENMTAGFYOSPOORVERLLKMTSKTPLGFSYDTRCF
DSVTEQDFVEBEIYQCNLDEFEARKVYISLLTERLYCGPMENSGACGVRRCRAS
GVLPSTFGNTICYIKATAAKAAGNPNDFLVCDDLVVAESDGVDRDTRALRAFT
EAMTRYGAPPDQAFQPTDYLELITSCSNVSVALDNKRYLYITKDATIPLARAWE
TARHTPVNRLHGLSIIMVAPTIWVWMTHTFFSILQSOILDRPLDFEMRYGATYSVP
LDLPAITERLHGLSFTLHSHSYSPVNLNRVAGTLRKLQCPPLRAWRHRAVRAKLIQ
GGRAKICGLYLEWAVRTKNLITPLPATGQLDLSWFTVGVGNDIYHVSVRARTYL
LLCLLLLVGVGIFLLPAR"

299. .871
/product="core protein"
872. .1447
/product="envelope protein"
1448. .2503
/product="NS1 protein"
2504. .3334
/product="NS2 protein"
3335. .5287
/product="NS3 protein"
5288. .5449
/product="NS4a protein"
5450. .6232
/product="NS4b protein"
6233. .7588
/product="NS5a protein"
7589. .9361
/product="NS5b protein"

BASE COUNT 1988 a 2647 c 2574 g 2181 t
ORIGIN

Query Match 93.0%; Score 585.6; DB 14; Length 9390;
Best Local Similarity 96.2%; Pred. No. 1.4e-152;
Matches 600; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 2 TGGGTAGGTGATCGATACCCCTTACGTGGCGATTCCGCCGATCTCATGGGGTACATCCCGC 61
DB 654 TGGGCAAGGTGATCGATACCCCTTACGTGGCGATTCCGCCGATCTCATGGGGTACATCCCGC 713

QY 62 TCGTCGGCGCTCCCGTAGGAGCGCTCGAAGAGCGCTTCGCGCATGCGCGTGGAGGCCCTTG 121
DB 714 TCGTCGGCGCTCCCGTAGGAGCGCTCGAAGAGCGCTTCGCGCATGCGCGTGGAGGCCCTTG 773
QY 122 AAGACGGGATAAATTTTCGCAACAGGGAATTTGCGCGGTTCCTCTTCTATTTTCCCTTC 181
DB 774 AAGACGGGATAAATTTTCGCAACAGGGAATTTGCGCGGTTCCTCTTCTATTTTCCCTTC 833
QY 182 TCGCTCTGTTCTTCTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGGAATACTGCTG 241
DB 834 TTGCTCTATTTCTTCTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGGAATACTGCTG 893
QY 242 GCCTCTATGCTTACCAACAGCTGTTTCCAATAGCAGTATTTGTGTACGAGGCGGATGACG 301
DB 894 GCTCTATGCTTACCAACAGCTGTTTCCAATAGCAGTATTTGTGTACGAGGCGGATGACG 953
QY 302 TTATTTCTGCAACACCGCGCTGATACCTTGTGTCTGAGCAGCGCAATACATCCACCTGCT 361
DB 954 TCATTTCTGCAACACCGCGCTGATACCTTGTGTCTGAGCAGCGCAATACATCCACCTGCT 1013
QY 362 GGACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACGCTTCGATAC 421
DB 1014 GGACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACGCTTCGATAC 1073
QY 422 GCAGTCTATGTGGACCTTATAGTGGCGCGGCCAGCAGTGTCTGCGCTCTACGTGGGTG 481
DB 1074 GCAGTCTATGTGGACCTTATAGTGGCGCGGCCAGCAGTGTCTGCGCTCTACGTGGGTG 1133
QY 482 ACATGTGTGGGCTGTCTTCTGCTGGGCAAGCCCTTACGCTTACAGCTCTGCGCATC 541
DB 1134 ATATGTGGGAGCGCTCTTCTGCTGGGCAAGCCCTTACGCTTACAGCTCTGCGCATC 1193
QY 542 AAACGGTTCAGACCTTGTAACCTGCTGCTGTAACCGGCGCATCTTTACGACATCGAATGG 601
DB 1194 AAACGGTTCAGACCTTGTAACCTGCTGCTGTAACCGGCGCATCTTTACGACATCGAATGG 1253
QY 602 CTGGGATATGATGATGAAATGGT 625
DB 1254 CTGGGATATGATGATGAAATGGT 1277
RESULT 8
HPCGSGS 9456 bp RNA linear VRL 04-FEB-1999
LOCUS Hepatitis C virus complete genome sequence.
DEFINITION D17763
ACCESSION D17763
VERSION D17763.1 GI:514395
KEYWORDS polyprotein; C; E1; E2/NS1; NS2; NS3; NS4a; NS4b; NS5a; NS5b.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 48)
AUTHORS Sakamoto,M., Akahane,Y., Tada,F., Tanaka,T., Woodfield,D.G. and Okamoto,H.
Entire nucleotide sequence and characterization of a hepatitis C virus of genotype V/3a
J. Gen. Virol. 75 (Pt 7), 1761-1768 (1994)
MEDLINE 94292926
PUBMED 8021605
REFERENCE 2 (sites)
AUTHORS Sakamoto,M.
JOURNAL Unpublished
AUTHORS Unpublished
REFERENCE 3 (bases 1 to 9456)
AUTHORS Okamoto,H.
Direct Submission
TITLE Submitted (27-SEP-1993) Hiroaki Okamoto, Jichi Medical School, Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi 329-04, Japan [E-mail:hokamoto@jichi.ac.jp, Tel:0285-44-2111(ex.3334), Fax:0285-44-1557]
JOURNAL
FEATURES Location/Qualifiers
source 1. .9456 /organism="Hepatitis C virus"

/mol_type="genomic RNA"
/isolate="NZL1"
/db_xref="taxon:11103"
1. .339
340. .9405
/codon_start=1
/product="polyprotein"
/protein_id="BAA04609.1"
/db_xref="GI:1183029"
/translation="NSTLPQPKTKRNTIRPQDVKPPGGQGVGVVYVLRPRRL
GVRATKRSQPRGRQPIPKARRSEGRSWAQYPMPLVYNGSCGWAHGLSPRG
SRSPGPNDRSRRLGRKVIDITCGPADLMGYIPLVGA PVGVARALAGVRALED
GNEATGNLPGCSFSIFLALFSLIHPAASLEWNTSLGLVLTNDNSSIYVZAD
VILHTPGVCPVQDQNTSCWTPTVAVRTVGAITASIRSHVLLVGAATVSCALY
VDMCGAVTFQGAFTFRRHQTQVOTCMCSLPGHLSGRHMAWMMNWSVAVGVV
AHVLPQFLFDIMAGAHWGIILAGLAYSMMQGNWAKVAIIMVMSGIADHTYTTGGTA
SRHTQAFAGLFDGQOQLQVNTNGSHWINTALNCNESINTGFIYFHKFNST
GCPQLSSCKPTTFRQMGPLTDANIITGSDRPVCHVYAPPCDIIPASSVCGPVY
CTPSPVVVGTTDARGVPTTWGENEKOVFLLKSORPPSGRWFSGSNWSTGFLTGC
APPCNIYGEGBPHNESDLFCPTDFRKHPTTYSRCGAGPWLTPRCWMDPYRLWHY
PCTVDFRLFKVPMFVGGFHRFTAAACNMTWGRGCDIEDRDRSEQLHSTTETELIIP
CSFTMPALSTGLIHLQNIIVDQYLYGVSGMGWALKWFEVILFVLLADARVCVA
LWMLMSOTEALENLVTLNAAAGTIGWLVAFCAWYVKGKLPVLTYSLTG
LWSLALLVLLPQAVANGSDSATLGAIVLFGFETLSPWYKMGRLMWNYOYI
CECEALHWVWPLIARGSRDGVILLTSLLYPSLIEDITKLLMVLGPLYLIQAITT
TPYVRAHVLRCLMVRVSGIKYFQMIILSIGRWENTLYDHLAPWQHWAAAGLKD
LAVATPFISSPMEIKVITWADTAAACDILCLGVPSARLREVLHGADDDYREMGWR
LLAPITAVTAQKRLGTLTSLTGRDKNVVTGEVQVLSATQTFGLTGVGVITWYV
HAGSRTLAGAKHPALOMYTNVDQDLVGPAPGAKELEPCACGSSDLVLTVDADVI
PARRGDSTASLLSRLPCLAKSGSGPVMPCHSHVAGI FRAAVCTRGVAKSLFIPV
ETLSQARSPPSNDSTPAVQSQVGLYHAPTSGSKTKVPAAYVAGQYVNLVNP
SVATLFGFSFMSRAYIDPNRNTVTAKLTYSTYKFLADGGCGGAYDVI
DECHQADATSLIGTGLVDQAEATAGRLTVLATATPGSITVPSHNEEVALGSEGE
IPYQKATP IALLKGRHLIFCHSKKCCDEIASKLGRMGLNAYVYRGLDVSIVPTTG
DVVATDALMTGDFDSDVDCNVAYQVDFSDPTFSIETRAPDADVRSORR
GRTGRGLGTAYVYASGRRPSCGVSVLCEYDAGCSWYDLQPAETVRLRAYLSTP
GLPVQDHLDFWESVFTGLTHIDHFLSQTKOQGLNFSVLTAYQATVCAQAQPPSW
DEWKKLVLKPTLHGTPPLLYRLGPVQNETCLTHPTIKYLMACSDALEVTSTWVL
LGGVLAALAAVCLSVGCVVIVGHIIELEGKPLVPDKEVLYQYDEMESCQAAPYIEQ
AQVIAHQFEKILGLQATQOQAVIEPIVTNNWQLEAFMHKHMWNVFSQYQLAGL
STLPGNPAVASLMAFASVTSPLTNTQMFENILAGWVATHLAGPQSSAFVVSGLAG
AAIGGILGRVLLDILAGVAGVSGALVAFKTMGCECPTAEDMNLIPALLSPGALV
GVICAALRHVVGEGAVQWNRLLIAPASRGNHVSPTHYPESDAARVATLLSLT
VTSLLRLHQWLNEDYPPSCDDMLRTIDWVCSVLADFKWLSAKIMPALPGLPFIIS
COGKYGVRGWDGVMSTRCPGAAITGHVKGMSRLAGPRTCANMWHGTFPINEYTTG
PSTPCPSNYTRALMRVAANSYVEVRRVGDHYITGATEDELCPCQVPAAEFFTEVD
GVRLHAYAPCKPLARDITTFWVGLHSYITIGSOLPECEPDVSVLTSLMRDPSHTAE
TAARLRGSPSEASSASQLSAPSLKATCQTHRPHPDAELVDANLLWRQBMGNIT
RVSETKVVVLDSPFLRATDDVPSVAECFKPPKTPPALPIIWARPDYNNPDLDR
WKAPDYVPTVHGKALPRGAPPVPPRRKRTIQLDGNVSAAALAAEKSPSPSKPQ
EENSQSGVDYDTSSTSKVPPSPGSESDSCSMPLLEGEPDPLDSCDSWSTVSDS
RQSVVCCSMYSWTKALITPCSABEELPI SPLNSLLRHHNLVYSTSRASQROK
KYTPRLOVLDHYKALKERASRVKARMLTIERACALVPHSARSKFCYSADV
RSLSRANQIRSVWEDLLEDTTPIPTIMAKNEVFCVDPAGKGRKPARLLIYEDLG
VVCSEKRALYDIQKLSITETMGPAFYQSPQORVERLLKMWTSKKTPLFGSYTRCF
DSTVTEQDRIIVEKLSYQCNLEPBAKVISSILTERLYCGGPMFNKQACQGYRRCAR
GVLPSTFGNTICYIKATAAANKLNPDLVCGDILVVAESDGVDEDRALRAFT
EAMTRVSPAPGAPQATYDLELITCSSNVSVARDKGRYYLTRDATTPLARAWE
TARHTPVNSMLGNIIMYARTINRWMMTHFESILQSQELDRPLDFEMYGATYSVTP
LDLPAIIEKHLGSLFTLHSYFVELNRVAGTLRKGLCPPLRAWRHRAVRARKLIAQ
GGAKICGLYLFNVAVRTKNTLPLPAAGQLDLSNFTVGVGNDRIYHSVSRARTRHL
LILCLLLLVGVGIFLLPAR"
340. .912
/product="C"
913. .1488
/product="E1"
1489. .2544
/product="E2/NS1"
2545. .3375
/product="NS2"
3376. .5328
/product="NS3"
5329. .5490

mat_peptide 5491. .6273
mat_peptide 6274. .7629
mat_peptide 7630. .9402
3'UTR 9406. .9425
polyA_site 9425
BASE COUNT 1998 a 2684 c 2578 g 2196 t
ORIGIN
Query Match 93.0%; Score 585.6; DB 14; Length 9456;
Best Local Similarity 96.2%; Pred. No. 1.4e-152;
Matches 600; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 2 TCGGTAAGGTCAATGATACCCCTTACGTGGGATTCGCGGATCTCATGGGGTACATCCCGC 61
DB 695 TGGGTAAAGTCAATGATACCCCTTAAAGTGGGATTTGCCGACCTCATGGGGTACATCCCGC 754
QY 62 TCGTCGGGCTCCCGTAGGAGCGGTGCAAGAGCCCTTTCGCGATGCGGTGAGGGGCCCTTG 121
DB 755 TCGTCGGGCTCCTGTAGGAGCGGTGCAAGAGCCCTTTCGCGATGCGGTGAGGGGCCCTTG 814
QY 122 AAGACGGGATAAATTTGCAACAGGAGATTTCCCGGTTGCTCTTTCTATTTTCTTTC 181
DB 815 AAGACGGGATAAATTTGCAACAGGAGATTTCCCGGTTGCTCTTTCTATTTTCTTTC 874
QY 182 TCGCTCTGTTCTTCTTAAATTCACGAGCAGTAGTCTAGAGTGGCGGGAATACGTCTG 241
DB 875 TTGCTCTGTTCTTCTTAAATTCACGAGCAGTAGTCTAGAGTGGCGGGAATACGTCTG 934
QY 242 GCCTCTATGTCCTTACCACGACTGTTTCAATAGCAGTATTTGTATGAGGCGCGATGAG 301
DB 935 GCCTCTATGTCCTTACCACGACTGTTTCAATAGCAGTATTTGTATGAGGCGCGATGAG 994
QY 302 TTATTCTGCACACACCGCGCTGCATACCTTGTGTTCAGACGCGCAATACATCACTGCT 361
DB 995 TCAATTCGCACACACCGCGCTGTGTACCTTGTGTTCAGACGCGCAATACATCACTGCT 1054
QY 362 GGACCCCACTGACACTCAGTGGGAGTCACTAGTGGGAGCAACACCGCTTCGATAC 421
DB 1055 GGACCCCACTGACACTCAGTGGGAGTCACTAGTGGGAGCAACACTACTGCTTCGATAC 1114
QY 422 GCAGTCATGTGACCTTATTAGTGGGCGCGCCACCATGCTGCTGCTCTTACGTGGGTG 481
DB 1115 GCAGTCATGTGACCTTATTAGTGGGCGCGCCACCATGCTGCTGCTCTTACGTGGGTG 1174
QY 482 ACATGTGGGGCTGTCTTCTTCGTGGGCAAGCCCTTCACTGAGACCTGCTGCGCCATC 541
DB 1175 ATATGTGGGGCTGTCTTCTTCGTGGGCAAGCCCTTCACTGAGACCTGCGACCCATC 1234
QY 542 AAACGGTCCAGACCTGTAACCTGCTGCTGACCCAGGCCATCTTTTCAGGACATCGAATCG 601
DB 1235 AAACGGTCCAGACCTGTAACCTGCTGCTGACCCAGGCCATCTTTTCAGGACATCGAATCG 1294
QY 602 CTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 625
DB 1295 CTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1318
RESULT 9
LOCUS HPCUS114CE
DEFINITION Hepatitis C virus genes for C, E and E2/NS1, partial cds.
ACCESSION D14309
VERSION D14309.1
KEYWORDS C; E; E2/NS1; core protein; envelope protein; nonstructural protein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; serNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

```
REFERENCE
AUTHORS      1 (sites)
              Okamoto,H., Tokita,H., Sakamoto,M., Horikita,M., Kojima,M.,
              Iizuka,H. and Mishiro,S.
TITLE        Characterization of the genomic sequence of type V (or 3a)
              hepatitis C virus isolates and PCR primers for specific detection
JOURNAL      J. Gen. Virol. 74 (Pt 11), 2385-2390 (1993)
MEDLINE      94065664
PUBMED       7504073
FEATURES     source
              2 (bases 1 to 1786)
              Okamoto,H.
              Unpublished
              Submitted (28-JAN-1993) to DDBJ by:
              Hiroaki Okamoto
              Immunology Division
              Jichi Medical School
              Kawachi
              Tochigi 329-04
              Japan
              Phone: 0285-44-2111x3334
              Fax: 0285-44-1557.
              Location/Qualifiers
                1..1786
                /organism="Hepatitis C virus"
                /mol_type="genomic RNA"
                /isolate="US114"
                /db_xref="taxon:11103"
                <1..277
                /citation=[1]
                278..>1786
                /codon_start=1
                /product="polyprotein precursor"
                /protein_id="BAA03255.1"
                /db_xref="GI:471113"
                /translation="MSTLPKPKRKNTIRPQDKVFPFGGQIVGGVYVLPRRGPRL
                GVPRATKTSRSPRRROPVPRKRSRSGSWAOGYPWPLYNREGCGWAGWLLSPRG
                SRFSWGNDRSRNRNLGKVDLTGCFADLMGYIPLVGAPVGGVARLAHGVRALED
                GINFATNGPGCSFSFLALSLCLHPASLEWNTSGLYVLTNDSCNSSTIVYBAD
                VILHTPGCTPCVDQGNVRCWTPVTPTVAVRVGATASIRSHVDLLVGAATWCSALY
                VGMCGAVFLVGAQTFRRPRRHQTVTCNSLYPGHLSGHRMADMMWNSPAYGMVY
                AHVLRIPQTLFDIVAGAHGILAGLAYISMQGNWAKVALIMVWFSGVDASTHTGSM
                ARDVYRTGFFARGPSQNLQLVNTGSHWLNRTALNCNESINTGFIAGLLYNNKFNST
                GCQRLSCKPITSFRQWGSLSLTDMANITDSSDDKPYCWHVAPRSCDTPASSV"
              mat_peptide 278..850
              /product="C"
              mat_peptide 851..1426
              /product="E"
              mat_peptide 1427..>1786
              /product="E2/NS1"
              BASE COUNT 353 a 522 c 507 g 404 t
              ORIGIN
              Query Match 92.4%; Score 582.4; DB 14; Length 1786;
              Best Local Similarity 95.8%; Pred. No. 9.7e-152;
              Matches 598; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
              QY 2 TGGGTAAGTCATCGATACCTTACGTGCGGATTTCGCGATCTCATGGGGTACATCCCGC 61
              |||||
              DB 633 TGGGTAAAGTCATCGATACCTTACGTGCGGATTTCGCGGACCTCATGGGGTACATCCCGC 692
              |||||
              QY 62 TCGTCGGCGCTCCCGTAGGAGCGCTGCGAAGCGCCTTGGCGCATGGCGTAGGGCCCTTG 121
              |||||
              DB 693 TCGTCGGCGCTCCCGTAGGAGCGCTGCGAAGCGCCTTGGCGCATGGCGTAGGGCCCTTG 752
              |||||
              QY 122 AAGACGGGATAAATTTCCGACACAGGGAATTTGCCGGTGTCTCTTTTCTATTTCCTTC 181
              |||||
              DB 753 AAGACGGGATAAATTTCCGACACAGGGAATTTGCCGGTGTCTCTTTTCTATTTCCTTC 812
              |||||
              QY 182 TCGCTCTGTTCTCTTCTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGCTG 241
              |||||
              DB 813 TTGCTCTGCTCTCTTCTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGCTG 872
              |||||
              QY 242 GCCTCTATGTCCTTACCAACGACTGTTCCCAATAGCAGTATTGTGTACGAGGCCGATGACG 301
              |||||
```

```
873 GCCTCTATGTCCTTACCAACGACTGTTCCCAATAGCAGCATTGTGTATGAGGCCGATGACG 932
302 TTATTCTGCACACACCCGGCTGCATACCTTGTCTCCAGACGGCAATACATCCAGTGTCT 361
|||
933 TTATTCTGCACACACCCGGCTGTATACCTTGTGTTCAGACGGCAATGTATCCAGGTGTCT 992
|||
362 GGACCCCACTGACACCTACAGTGGCAGTCAAGTACGTTCGGAGCAACACCCGCTTCGATAC 421
|||
993 GGACCCCACTGACACCTACAGTGGCAGTCAAGTACGTTCGGAGCAACACCCGCTTCGATAC 1052
|||
422 GCAGTCATGTGGACCTATTATTAGTGGGGCGGCGCACAGTGTCTCTGGCTCTACGTGGGTG 481
|||
1053 GCAGTCATGTGGACCTATTATTAGTGGGGCGGCGCACAGTGTCTCTGGCTCTACGTGGGTG 1112
|||
482 ACATGTGTGGGGCTGTCTTCCTCGTGGGACAGACCTTTCAGCTTCAGACCTCTCGCCATC 541
|||
1113 ATATGTGTGGGGCGGTCTTCTCGTGGGACAGACCTTTCAGCTTCAGACCTCTCGCCATC 1172
|||
542 AAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCGCATCTTTCAGGACATCGAATGG 601
|||
1173 AAACGGTCCAGACCTGTAACTGCTGTACCCAGGCGCATCTTTCAGGACATCGAATGG 1232
|||
602 CTTGGGATATGATGATGAATGGT 525
|||
1233 CTTGGGATATGATGATGAATGGT 1256
|||
```

```
RESULT 10
LOCUS      HPCTH85CE 1786 bp RNA linear VRL 01-FEB-2000
DEFINITION Hepatitis C virus genes for C, E and E2/NS1, partial cds.
ACCESSION D14307
VERSION    D14307.1 GI:456468
KEYWORDS   C; E; E2/NS1; core protein; envelope protein; nonstructural
           protein
SOURCE      Hepatitis C virus
ORGANISM    Hepatitis C virus
REFERENCE   1 (sites)
AUTHORS     Okamoto,H., Tokita,H., Sakamoto,M., Horikita,M., Kojima,M.,
           Iizuka,H. and Mishiro,S.
TITLE       Characterization of the genomic sequence of type V (or 3a)
           hepatitis C virus isolates and PCR primers for specific detection
JOURNAL     J. Gen. Virol. 74 (Pt 11), 2385-2390 (1993)
MEDLINE     94065664
PUBMED      7504073
REFERENCE   2 (bases 1 to 1786)
AUTHORS     Okamoto,H.
JOURNAL     Unpublished
COMMENT     Submitted (28-JAN-1993) to DDBJ by:
           Hiroaki Okamoto
           Immunology Division
           Jichi Medical School
           Kawachi
           Tochigi 329-04
           Japan
           Phone: 0285-44-2111x3334
           Fax: 0285-44-1557.
           Location/Qualifiers
             1..1786
             /organism="Hepatitis C virus"
             /mol_type="genomic RNA"
             /isolate="Th85"
             /db_xref="taxon:11103"
             <1..277
             /citation=[1]
             278..>1786
             /codon_start=1
             /product="polyprotein precursor"
             /protein_id="BAA03253.1"
             /db_xref="GI:471118"
             /translation="MSTLPKPKRKNTIRPQDKVFPFGGQIVGGVYVLPRRGPRL
```


GVRATKTSERSQPRGRQPIPKARRSRSRWAQPGYWPPLYGNEGCGWAGWLLSPRR SRPSPGPNDRRRSRLGKVIDTLCGFADLMGYIPLVGAAPVGGVARALAGHVRALD GINATFATPCPSFISILLALFSLIPHAASLEWNTSGLYLINDCSNSIYVEAD VILHTPGICPCVHNGNTSCWPTPTVAHVHVGATTASIRSHVDLLVGAATMCSALY VGDVCGAVFLVQAFITGPRRHOTVQTCNSLYPGHLSGHRMAMMMNWSFAGMVV AHLVRLPOTLFDIIAGAHWILAGLAYYSMOGNWAKVAIIMVPSGVAETRIITGTA AHDVVRPASLFRNARONLQLVNTGSHINRTALNCNDSINTGFIARLIFYHKNST GCPQLRSCKPITSFNQSGWGLTDENITGPRSDDKPYCHWYHAPRCKSVPASSV" 278...850 /product="C" mat_peptide 376 a 526 c 490 g 394 t 851...1426 /product="E" mat_peptide 1427...>1786 /product="E2/NS1"									
BASE COUNT 376 a 526 c 490 g 394 t									
ORIGIN									
Query Match 91.4%; Score 576; DB 14; Length 1786; Best Local Similarity 95.2%; Pred. No. 6e-150; Matches 594; Conservative 0; Mismatches 30; Indels 0; Gaps 0;									
Qy	2	TGGGTAAGTCAATGATACCTTACGTCGGGATTCGCCGATCTCATGGGGTATACATCCCGC	61						
Db	633	TGGGTAAGTCAATGATACCTTACGTCGGGATTCGCCGATCTCATGGGGTATACATCCCGC	692						
Qy	62	TCGTGGGCTCCCGTAGAGCGCTGCGAGAGCCCTTGGCGATGGCGTGAGGGCCCTTG	121						
Db	693	TCGTGGGCTCCCGTAGAGCGCTGCGAGAGCCCTTGGCGATGGCGTGAGGGCCCTTG	752						
Qy	122	AAGACGGGATAAATTCGCAACAGGGAATTCGCCGGTGTGCTCCCTTTCTATTTTCCCTC	181						
Db	753	AAGACGGGATAAATTCGCAACAGGGAATTCGCCGGTGTGCTCCCTTTCTATTTTCCCTC	812						
Qy	182	TCGCTCTGTTCTCTTGCTTAATTCATCCAGCAGCTAGCTAGAGTGGCGGAATACGTCGT	241						
Db	813	TTGCTCTGTTCTCTTGCTTAATTCATCCAGCAGCTAGCTAGAGTGGCGGAATACGTCGT	872						
Qy	242	GCCTCTATGCTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAGC	301						
Db	873	GCCTCTATGCTCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTATAGGCGCGATGAGC	932						
Qy	302	TTATTTCTGCACACCCGGCTGCATACCTTGTGTCAGGACGGCAATACATCAGCTGCT	361						
Db	933	TTATTTTGCACACACCCGGCTGTATACCTTGTGTTTCATAACGGCAATACATCAACGCTGT	992						
Qy	362	GGACCCAGTACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATAC	421						
Db	993	GGACCCAGTACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATAC	1052						
Qy	422	GCAGTCATGTGGACCTATTAGTGGCGCGGCCACAGTGTCTCTGCGCTCTACGTGGGTG	481						
Db	1053	GCAGTCATGTGGACCTATTAGTGGCGCGGCCACAGTGTCTCTGCGCTCTACGTGGGTG	1112						
Qy	482	ACATGTGTGGGCTGCTTCTCTCGTGGGACAGCCTTCAAGTTCAGACCTCTGTCGCCATC	541						
Db	1113	ATGTGTGTGGGCGCTCTTCTCTGCGGACAAAGCCTTCAGGTTTCAGGCCCTCTGCGCCATC	1172						
Qy	542	AAACGCTCCAGACCTGTAACCTGCTGCTGTACCCAGGCCATCTTTCAGGACATCGAATGG	601						
Db	1173	AAACAGTCCAGACCTGTAACCTGCTGCTGTACCCAGGCCATCTTTCAGGACATCGAATGG	1232						
Qy	602	CTTGGGATATGATGATGAATCGT 625							
Db	1233	CTTGGGATATGATGATGAATCGT 1256							
RESULT 11									
AF046866									
LOCUS									
DEFINITION Hepatitis C virus type 3a (CB), complete genome.									
ACCESSION AF046866									
VERSION AF046866.1 GI:2895898									
KEYWORDS									

SOURCE ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1 (bases 1 to 9425)
Shukla,D.D., Chaturvedi,S., Cao,J.Y. and Hoynes,P.A.
Complete Nucleotide Sequence of the genome of Hepatitis C Virus
type 3a (CB)
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 9425)
Shukla,D.D., Chaturvedi,S., Cao,J.Y. and Hoynes,P.A.
Direct Submission
TITLE
JOURNAL
Submitted (04-FEB-1998) Biomolecular Research Institute, 343, Royal
Parade, Parkville, Melbourne, Victoria 3052, Australia
FEATURES
Location/Qualifiers
1..9425
source

CDS

/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="type 3a (CB)"
/db_xref="taxon:11103"
340..9405
/codon_start=1
/product="polyprotein"
/protein_id="AA03058.1"
/db_xref="GI:2895898"
/translation="MSTLPKPKQTKRNTIRPDQVDFGGQIVGGVYVLPRRGRRL
GVCATKTSERSQPRRRQPIPKARRSRSRWAQPGYWPPLYGNEGCGWAGWLLSPRG
SRPSPGPNDRRRSRLGKVIDTLCGFADLMGYIPLVGAAPVGGVARALAGHVRALD
GINATFATPCPSFISILLALFSLIPHAASLEWNTSGLYLINDCSNSIVCBAD
VILHTPGICPCVHNGNTSCWPTPTVAHVHVGATTASIRSHVDLLVGAATMCSALY
VGDVCGAVFLVQAFITGPRHOTVQTCNSLYPGHLSGHRMAMMMNWSFAGMVV
AHLVRLPOTLFDIIAGAHWILAGLAYYSMOGNWAKVAIIMVPSGVAETRIITGTA
AHDVVRPASLFRNARONLQLVNTGSHINRTALNCNDSINTGFIARLIFYHKNST
GCPQLRSCKPITSFNQSGWGLTDENITGPRSDDKFYCHYHAPRCKSVPASSV"
PCTVNTFLKVRMFGEHFRPTAACNMTGERCDIEDRSEQLHLHSTTELAILP
CSFTMPALSTGLIHLQNVIVQYLVGGSMGVALKWFEVILI FLLADRRVCVA
LWLMIMTQABAALENLTLNAAAGTHGIGWLVAFCAAWYRGKLVVTVYSLTG
LWPLAILVLMQRAYAWSGDSATLGAGVLALFGFTLS PMWIKGRLIWNWVITI
CRESALOVWVPLLLARGSRDGVLLTSLYPSLIFDITKLILVGLYLILQAQVTT
TPYFVRQVLRCLMVRSMVGKGYFQMLISIGRNFNTLYDHLAPWQWAAAGLKD
LAVATPYFSPMEIKVITWGAADTAACDILGCLPVSARLREVLLGPAADDYREMGWR
LLAPATVIAQQTGRLGTLTSLTGRDNVTVGEVLSLTATQTFLGTGGVGTWTV
HGASRTLAGAKHPALOMTVNDQLVGMPAPPGAKSLPCCAGSADLYLVTRDADVI
PARREDSTASLLSPRLACLKSGSGPVMCPSGHVAGIFRAAVCTRGVAKALQIPV
ETLSQTSPSPSDNSTPAPVQSYQVGYLHAPTSGSKSTKVPAYVAGQYTVLNLN
SVATLGFSGFMSRAYGIDPNIRNTVTGAKLTISTYKFLADGGSGGAYDVI
CDEHAQDATSLIGLTVLQDAETAGVLTATLTPGSIITVPHSNTEEVALLSGE
IPFYKATIPAILKGRHLIFCHSKKCDDEIASKLRGMGLNAYAYRGIDDSVIPTTG
DVVVCATALMTGTGDFSDVIDCNVAVQVDFSLDPTFSIETRTAPDASRSORR
GRTGRGLGTVRYVAPGRSPGMPDVSILCEYDAGCSWYDLOQAEVTVRLAYLSTP
GLPVCQDHLDFWESVFTGLTHDAHLFSQTKQGLNFSLYLTAYQATVCARAQAPPSW
DETWKLVRLLKPTLHPTLPLVRLGFPVQNEICLSHPTIKYNMACSDALEVITSTWV
LGGVLAALAAVCLSVGCVVIIGHI BLGGKPAVLPDKVLYQQYDEMEESQAPYIEQ
AQATLAHQEKVGLQORATQOQAVIEPTVATNMOKLEAFWKHMWVFSQYVLAGL
STLPCNPAVSLMAFTASVTSPLTNTMFFNIGLVATHLAGPQSSAFVSVGLAG
RAIGIGIGVRLVLLDLAGVAGSALVAFKIMGSELPTAEDMNLPAILLSPGALV
GVI CHAIRRHVGPGEAGQWNRLLI AFASRGNHVSPHYVPESDAAARVALLSLT
VSKLRLRHQWGINEDYPSDCSDWLRI IWDVWCVSLSDFKSWLSAKIMHGLPGLPFI
COKGYKMGWGDVMSRCPGATITGHVXGMSRLAGPRTCANMWHGTFPFPNFTTG
PSTPCSPNYTRALMVAANSYVVRVGDPHYITGATEDELKCPQVAPAEFPFTEVD
GVRLHRYAPCKPLAREBITFVSGLNSVTIGSQLCEPEPDVSLTSMRLDPSHTAE
TAARELAGSPSEASSASQLSALSKATCQTHRPDPAELVNLNLRWQMGSNIT
VRESKTVILDSPEPLRAETDAAESVAABCFKKPKYPPALPIWARPDYNPPLDOR
WKADYVPTVHGCALPRGAPPVPPRRKRTIQDGSNVSAALAEKSPSLKPKQ
BENSSSGVDTSQSTTSKVPSPGSDSES CSNMPLEGEPEGDPDLS CDSWSTSDS
EEQSVVCSMSYMTGALITPCSABEELPISPLNSLLRHNLVYSTSRSASORQR
KVTDFRLOVLDHDKYTVLKEVKERASRVKARMLTIEACALVPHPSAKFCYSKDV
RSLSSRAINQIRSWEDELLEDITPTTITMAKNVEFVCDVPAKGRKPARLIVYDFLG
VRVCEKRALYDVIQKLALETIGSAVQFQISPOORVERLLKMTSKTTFGFSYTRCF
DSTVTEQDIRVEETIYQCNELEPEARKVISLTERLYCGPMFNKSGACGRCRRAS
GVLPTSGNTITCYIKATAAANGAGLRDPLVLCVGDLLVVAESDGVDEBDGAALRAFT

BASE COUNT		1979 a	2699 c	2599 g	2148 t
ORIGIN		BAMTRYAPPDGAQPTVDLELITSCSSNVSVARDDKGRYYVLTTRDATTPLARAAME TARHTPVSNGNLIIMAPTIWVRVMTWTFESLOSQELIDRLPDEMFGATVYVTP LDLPAIIBERLHGLSAGTLLHSTVPELNRVAGTLAGKCPPLRARHRAARAKLIAQ GKAKICGLYLFNVAWTKTKTLPFRAGQDLISWFTVGVGGNDIYHVSARATRYL LCLLLLTGVGIFLLPAR"			
Query Match		91.4%;	Score 576;	DB 14;	Length 9425;
Best Local Similarity		95.2%;	Pred. No. 6.6e-150;		
Matches 594;		Conservative 0;	Mismatches 30;	Indels 0;	Gaps 0;
QY	2	TCGGTAAGTTCATGCATACCCCTTACGTCGGATTCGCCGATCTCATGGGGTACATCCCGC	61		
DB	695	TGGGTAAGTTCATGCATACCCCTTACGTCGGATTCGCCGATTCATGGGGTACATCCCGC	754		
QY	62	TGCTGGCGCTCCCGTAGAGCGCTGCGAAGAGCCCTTCGCGATCGGTGAGGGCCCTTG	121		
DB	755	TCGTGGCGCCCGCTGGGAGCGCTGCGAAGAGCCCTTCGCGATCGGTGAGGGCCCTTG	814		
QY	122	AGACGGGATAAATTCGCAACAGGGAATTCGCCGTTGCTCTTCTTCTATTTCTCTTC	181		
DB	815	AGACGGGATAAATTCGCGACAGGGAATTCGCCGTTGCTCTTCTTCTATTTCTCTTC	874		
QY	182	TCGCTCTGTTCTTCTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC	241		
DB	875	TTGCTCTGTTCTTCTCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTC	934		
QY	242	GCTCTATGTCCTTACCAACGACTGTTCATATGAGTATGTTGTCAGGAGCGGATGAGC	301		
DB	935	GCTCTACGTCCTTACCAACGACTGTTCATATGAGTATGTTGTCAGGAGCGGATGAGC	994		
QY	302	TTATTCTGCACACACCGGCTGCATACCTTGTCTCCAGGACGCAATACATCCAGTGT	361		
DB	995	TCATTCTGCACACACCGGCTGGTACCTTGTCTCAGAACGACAATATATCCAGTGT	1054		
QY	362	GGACCCAGTGACACTTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATAC	421		
DB	1055	GGACCCAGTGACACTTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATAC	1114		
QY	422	GCAGTATGGGACCTTATAGTGGCGCGCAGCAGTGTCTCTGCGCTCTACGTGGGTG	481		
DB	1115	GATGATGATGGACCTTATAGTGGCGCGCAGCAGTGTCTCTGCGCTCTACGTGGGTG	1174		
QY	482	ACATGTGTGGGCTGTCTTCTCTGTGGGAACAGCTTTCAGTTCAGACCTCTCGCCATC	541		
DB	1175	ATATGTGTGGGCTGTCTTCTTGTGGGAACAGCTTTCAGTTCAGACCTCTCGCCATC	1234		
QY	542	AAACGGTCCAGCCTGTAAGTCTGCTGTGTACCCAGGCGCATCTTTACAGGACATCGAATGG	601		
DB	1235	AAACAGTCCAGCCTGTAAGTCTGCTGTGTACCCAGGCGCATCTTTACAGGACATCGAATGG	1294		
QY	602	CTTGGGATATGATGATGAATGGT	625		
DB	1295	CTTGGGATATGATGATGAATGGT	1318		
RESULT 12		HPCK3A 9454 bp RNA linear VRL 07-FEB-1999			
HPCK3A		Hepatitis C virus complete genome sequence.			
DEFINITION		Hepatitis C virus complete genome sequence.			
ACCESSION		D28917			
VERSION		D28917.1 GI:558520			
KEYWORDS		polyprotein.			
SOURCE		Hepatitis C virus			
ORGANISM		Hepatitis C virus			
REFERENCE		Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
AUTHORS		Yamada,N., Tanihara,K., Mizokami,M., Ohba,K., Takada,A., Tsutsumi,M. and Date,T.			
TITLE		Full-length sequence of the genome of hepatitis C virus type 3a: comparative study with different genotypes			

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

J. Gen. Virol. 75 (Pt 11), 3279-3284 (1994)
95053917
7964640
2 (bases 1 to 9454)
Date, T.
Direct Submission
Submitted (12-MAR-1994) Takayasu Date, Kanazawa Medical University,
Department of Biochemistry; Uchinada, Kahoku-gun, Ishikawa 920-02,
Japan (Tel:0762-86-2211(ex.3701), Fax:0762-86-4693)
Location/Qualifiers
1. .9454
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="K3a"
/isolate="HCV-K3a/650"
/db_xref="taxon:11103"
340. .9405
/codon_start=1
/product="polyprotein"
/protein_id="BAA06044.1"
/db_xref="GI:558521"
/translation="MSTLPKPKRRTKRTIRPQDVKFGGGVYVGVVLPKRRGL
GVATRTKTSRSPGRRRKPIPKARSEGRSNAQPGVPMPLVGNCEGAGWALLSPRG
SRFNWPNDRPRRSNLGKVIDLTCTGADLMGYIPLVGAPLGGARALAHVRALED
GINFATNPGCSFSIFLLALFSLIHPAASLEWRNTSGLYLTNDCSNSSIVYEADD
VILHTPGCI PCVQDNTSTCTPTVTVAVRYGATTASIRSHVDLLVAGAGTWCALY
VDMCGPVLVQDFTFPRRRHTVQTCNSLYPHLSGQRMADMMNWSVAVGQA
AHILRPLQTLFDVAGAHGCI IAGLAYSMQGNWAKVAIIMVFSGVDAVTHVTAQA
ARNAYGISLFSVGAQNQLINTSGNSHINRTALNCNESINTGTAGLTFYHKFNS
GCPQRUSCKPTTFKQGGWGLTDANTGSDDKPYCWHYAPRCPGIVPALNVCOPY
CTFPVYVGTDAKGAFTYTGANKTDVFLLESAPTPPSPGSGWFCCTWMNSTGFVLTG
APTCNIYGDQDQANESDFCTDCFRKHPEATYSRCCGAGFWLTPLCLVDIYPRLWHY
PNTFTLFKVRMFVGGFHRFAACNWTGRGECDEIDRDSBOHPLHLHSTIELALIP
CSFTPMALSTGLIHLHONIVDQVLYGIGSGMVGWALKWFVILI LLLADLADLCA
LMLIITISQAEALENLAVNAAGTGHGVLVAFCAAWYRGKLVPLVYLTG
LMSALHVLPLPQRAYNSGDSATLGGILVLPFTLSPYWKHMLARLIWNQYTI
CRCSALHVVPPPLLAGRGDGVILLTSLYPSLI FDI TKLLAALGPLVLIQATTA
TPYFRAHLVRLCMLVRSVMGKIFQMI LLSLADGNTYLYDHLAPMOWAAAGLKD
LAVATEPVI FSPMEIKVITWAGDAACTLCGLPFSARLGRVLELLGAPDADYREMR
LLAPITAYAGVQTRGLGTIVTSLGRDNVAGVEVLSLTATQTFGLTIVGVNWMVYI
HGASRTLAGVKHPALQYNTVDODLVGMPAPPAKSLFECTCGSADLYLVTRDADVI
PARRGUSTASLLSPRLARKSGSGGPPCPGSHVAGIKPRAAVCTRGVAKALQFI PV
ETLSQAKSPFSNTPFPAYQSYQVYLHAPGTGSKSTKVPAAVYVQAGQVNLVPL
SVAATLPGFSFMSRAYGDPNIRTNRTVTGAKLTYSTYKGLAGGCGCGADVII
CDDCHAODATSLIGTIVLQDQETAGVRLTVKATKATGPNISVAYYRGLDVSVPITTC
IPFYGATPIACIKIGRHILI FCHSKKCDKMGMLNGLNAVAYYRGLDVSVPITTC
DVTGRLGTVYVTPGPRSGMDFSVVLCYDAGCSWYDLQYATVTCARQAQPPSW
GLPVCQDHLDLWESVFTGLTHIDAHFLSOTQAGLNFSLYATQATVTCARQAQPPSW
DGTWCLVRLKPTLHGPTLLYRLGPVONEICLTHPI TKYVVMACMSADLEVTITWVL
LGGVLAARAAVCLSVGCVVI VGHIELGKPAVDPKVLVYQYDMECSQARPIEQ
AQVIAHQFKEKYLGLLQATQOAVIEPI VVSNWQKLEVLNKHKNWNSFGIQLYLAG
STLPGNPAVALMAFTASVTPSLTNTNMFNLLGGWVATHLAGQAASA FVVSGLAG
AAIGIGLRLVLLDILAGVAGSALVAFKIMGGEPTTEDMNLNLLPAILSPGALV
GVI CAAILRRHVGPCEGPVQWNRLLI AFASNGHVSAPHYVPEDAARATALLSLLT
VTSLLRLHWNEDYPSGCDGLRII IWDVCSVSVDFKTLWSAKI MPALPGLPFI S
CQKYGKVRWGDSVMSVTRCPCGASLHGVKSGMELAGPRTCANMCHGTFPINEYTTG
PSTPCPPNRYTRALMRVANSVYVRRVGDHYITGATEDGLKCPQVPAFEFFTEVD
GVRIHRYAPCPRLRLDEITFMVGLNSVAGSQLPCEPEPDSVLTSLMRDSSHITAE
TAARLARGSPSEASASQSLATCATQTHPHDPAELVNDANLLMRQMGSNIT
RYSESTVILDSFPLRAETDDELSAAECFKPKPKYPALP IWARDPNYPLLRD
WKSPPYVPTVHGCALPKGAPVPPPRKRTIQLDGSNVSAALAAKSPFSSPKQ
RSLSKAIIQIRSVWEDLEDITPTIPIIMAKNEVFCVDPKAGKKAARLLYDPLG
KVTDFRLQVLDHDKYKALQEVKERSKVRARMLSTEEACALVPHGKARSKFGVSAKD
RSLSKAIIQIRSVWEDLEDITPTIPIIMAKNEVFCVDPKAGKKAARLLYDPLG
VRYCEKALYDVIQIRLSITMGSAFGYSPQRVERLLKMMWTSKTKPIGFSYDFCF
DSVTFQDITRVEAVYQCCNLEPQQAIISSILTERLYCGPMNNSGAGCSDYDCRAS
GVLPTSGQDITCYIKATAARAAGNPDFLVGGDVLVVAESDGVGDDRTDRAFT
EAMTRYSAPGDAPOPTDLELTSCSSNVSVARDKGRYYVLTTRDATTPLARAAME
TARHTPVSNGNLIIMAPTIWVRVMTWTFESLOSQELIDRLPDEMFGATVYVTP
LDLPAIIBERLHGLSAGTLLHSTVPELNRVAGTLAGKCPPLRARHRAARAKLIAQ
GKAKICGLYLFNVAWTKTKTLPFRAGQDLISWFTVGVGGNDIYHVSARATRYL
LCLLLLTGVGIFLLPAR"

CDS

BASE COUNT	1981 a	2679 c	2601 g	2193 t	
ORIGIN	LILLLLLTGVGIFLLPAR"				
Query Match	91.4%; Score 576; DB 14; Length 9454;				
Best Local Similarity	95.2%; Pred. No. 6.6e-150;				
Matches 594; Conservative	0; Mismatches 30; Indels 0; Gaps 0;				
QY	2	TGGGTAAGTTCATCGATACCTTACGTGCGGATTTGCGCGATCTCATGCGGTACATCCCGC	61		
Db	695	TGGGTAAGTTCATCGATACCTTACGTGCGGATTTGCGCGATCTCATGCGGTACATCCCGC	754		
QY	62	TGCTCGGGCTCCCGTAGGAGCGCTGCGAAGAGCCTTCCGATGCGGTGAGGGCCCTTG	121		
Db	755	TGCTCGGGCTCCCGTAGGAGCGCTGCGAAGAGCCTTCCGATGCGGTGAGGGCCCTTG	814		
QY	122	AAGACGGGATAAATTTGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATTTTCCCTTC	181		
Db	815	AAGACGGGATAAATTTGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATTTTCCCTTC	874		
QY	182	TGCTCTGTTCTTCTGTTAATTCATCAGAGCTAGTCTAGAGTGGCGGAATACGTCTG	241		
Db	875	TTGCTCTGTTCTTCTGTTAATTCATCAGAGCTAGTCTAGAGTGGCGGAATACGTCTG	934		
QY	242	GCCTCTATGCTCTTACCAACGACTGTTCCAATAGCAGTATTGTTACGAGCGCGATGACG	301		
Db	935	GCCTCTATGCTCTTACCAACGACTGTTCCAATAGCAGTATTGTTACGAGCGCGATGACG	994		
QY	302	TTATTCTGCACACACCCGGCTCATACCTTGTGTCAGGACCGCAATACATCCAGTGT	361		
Db	995	TTATTCTACACACACCCGGCTGTATACCTTGTGTTGTTGAGGACCGCAATACATCCAGTGT	1054		
QY	362	GGACCCCAAGTACACTACAGTGGCAGTCAAGTACGTGCGAGCAACCCGTTTCGATAC	421		
Db	1055	GGACCCCAAGTACACTACAGTGGCAGTCAAGTACGTGCGAGCAACCCGTTTCGATTC	1114		
QY	422	GCAGTCATGTGACCTATTAGTGGCGCGGCGCACAGTGTCTGCGCTCTACGTGGGTG	481		
Db	1115	GCAGTCATGTGACCTATTAGTGGCGCGGCGCACAGTGTCTGCGCTCTACGTGGGTG	1174		
QY	482	ACATGTGTGGGGCTGTCTTCCTCGTGGGACAGCCTTCAGCTTCAGACCTGTCGCCATC	541		
Db	1175	ATATGTGTGGGGCGGTCTTCCTCGTGGGACAGCCTTCAGCTTCAGACCTGTCGCCATC	1234		
QY	542	AAACGGTCCAGACCTGTAACTGCTGCTGATCCCGGCCATCTTTTCAGGACATCGAATGG	601		
Db	1235	GGACGGTCCAGACCTGTAACTGCTGCTGATCCCGGCCATCTTTTCAGGACAGCGAATGG	1294		
QY	602	CTTGGGATGATGATGAATCGT	625		
Db	1295	CTTGGGATGATGATGAATCGT	1318		
RESULT 13	HPCHEM26CE				
LOCUS	Hepatitis C virus genes for C, E and E2/NS1, partial cds.				
DEFINITION	Hepatitis C virus genes for C, E and E2/NS1, partial cds.				
ACCESSION	D14311				
VERSION	D14311.1				
KEYWORDS	C; E; E2/NS1; core protein; envelope protein; nonstructural protein.				
SOURCE	Hepatitis C virus				
ORGANISM	Hepatitis C virus				
REFERENCE	Hepatitis C virus				
AUTHORS	Hepatitis C virus				
TITLE	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				
JOURNAL	Okamoto, H., Tokita, H., Sakamoto, M., Horikita, M., Kojima, M., Iizuka, H. and Mishiro, S.				
MEDLINE	Characterization of the genomic sequence of type V (or 3a) hepatitis C virus isolates and PCR primers for specific detection				
PUBMED	J. Gen. Virol. 74 (Pt 11), 2385-2390 (1993)				
	9405664				
	7504073				

REFERENCE	2 (bases 1 to 1786)				
AUTHORS	Okamoto, H.				
JOURNAL	Unpublished				
COMMENT	Submitted (28-JAN-1993) to DDBJ by: Hiroaki Okamoto Immunology Division Jichi Medical School Kawachi Tochigi Japan Phone: 0285-44-1557 Fax: 0285-44-2111x3334 Location/Qualifiers 1. .1786 /organism="Hepatitis C virus" /mol_type="genomic RNA" /iso1ate="HEM26" /db_xref="taxon:11103" 41. .277 /citation=[1] 278. .>1786 /codon_start=1 /product="polyprotein precursor" /protein_id="BAA03257.1" /db_xref="GI:471113"				
FEATURES	source				
5'UTR	/translation="MSTLPKPKRKTKRNTIRRPQDIKPPGGQIVGGVYVLPRRPRL GYRATKTSERSQPRRQPIPKARRSEGRSWAQFQWPLYGNCGWAGWLLSPRG SRPTGPNDRRRSRNLKVIDLTTCGADLMGYIPLVAGPVGGVARALAHGVRALED GINPFTNLPGCSFSIFILLALFSLIHPAASLEWRTNSGLYVLVNDSCNSSIYVBADD VILHPGCVPCVQDGNKSTCTPTPTVAVRVGATTASIRSHVDLLVGAATCSALY VDMCGAVLPGQAFTFRRHQVTQCNCSLYPGHLSGHQWANDMMNWSFAVGNV AHVLPQTLFDIIAGAHWGLIAGLYYSMQGNWAKVAIIIVMFSGVDABTYTGGSV AHATGIIASLFTVGARQKLQLVNTGSHWINSALNCNESINTGTAGLIFYHKFNST GCPHRLSSCKPITSFROGWSLTDANISSSSBDKPYCWYAPRPCTVVPASSV"				
mat_peptide	278. .850	/product="C"			
mat_peptide	851. .1426	/product="E"			
mat_peptide	1427. .>1786	/product="E2/NS1"			
BASE COUNT	363 a	517 c	504 g	402 t	
ORIGIN	Query Match 91.2%; Score 574.4; DB 14; Length 1786; Best Local Similarity 95.0%; Pred. No. 1.7e-149; Matches 593; Conservative 0; Mismatches 31; Indels 0; Gaps 0;				
QY	2	TGGGTAAGTTCATCGATACCTTACGTGCGGATTTGCGCGATCTCATGCGGTACATCCCGC	61		
Db	633	TGGGTAAGTTCATCGATACCTTACGTGCGGATTTGCGCGATCTCATGCGGTACATCCCGC	692		
QY	62	TGCTCGGGCTCCCGTAGGAGCGCTGCGAAGAGCCTTCCGATGCGGTGAGGGCCCTTG	121		
Db	693	TGCTCGGGCTCCCGTAGGAGCGCTGCGAAGAGCCTTCCGATGCGGTGAGGGCCCTTG	752		
QY	122	AAGACGGGATAAATTTGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATTTTCCCTTC	181		
Db	753	AAGACGGGATAAATTTGCAACAGGGAATTTGCCCGGTTGCTCTTTCTATTTTCCCTTC	812		
QY	182	TGCTCTGTTCTTCTGTTAATTCATCAGAGCTAGTCTAGAGTGGCGGAATACGTCTG	241		
Db	813	TTGCTCTGTTCTTCTGTTAATTCATCAGAGCTAGTCTAGAGTGGCGGAACACGTCTG	872		
QY	242	GCCTCTATGCTCTTACCAACGACTGTTCCAATAGCAGTATTGTTACGAGCGCGATGACG	301		
Db	873	GCCTCTATGCTCTTACCAACGACTGTTCCAATAGCAGTATTGTTACGAGCGCGATGACG	932		
QY	302	TTATTCTGCACACACCCGGCTCATACCTTGTGTCAGGACCGCAATACATCCAGTGT	361		
Db	933	TCATTCTGCACACACCCGGCTGTGCGCTTGTGTTCCAGATGGCAATAAATCCAGTGT	992		
QY	362	GGACCCCAAGTACACTACAGTGGCAGTCAAGTACGTGCGAGCAACCCGTTTCGATAC	421		

Db	993	GGACTCCAGTGACACCTACAGTGGCAGGTCAGGTATGTCGGAGCAACCAACGGCTTCGATAC	1052
Qy	422	GCAGTCATGTGGACCTATTAGTGGCGCGGCCACGATGTGCTCTGGCTCTACGTGGGTG	481
Db	1053	GCAGTCATGTGGACCTATTAGTGGCGCGGCTACGACGTGCTCTGGCTCTACGTGGGTG	1112
Qy	482	ACATGTGTGGGCTGCTCTTCCTCGTGGGACAAAGCCTTCACGTTACAGACCTCGTGGCATC	541
Db	1113	ATATGTGTGGGCTGCTCTTCCTCGTGGGACAAAGCCTTCACGTTACAGACCTCGTGGCATC	1172
Qy	542	AAACGGTCCAGACCTGTAACTGCTCGCTGTATCCAGGCGCATCTTTTACGAGCATCGAATGG	601
Db	1173	AAACGGTCCAGACCTGTAACTGCTCACTGTATCCAGGCGCATCTTTTACGAGCATCAATGG	1232
Qy	602	CTTGGGATATGATGAATCGT	625
Db	1233	CTTGGGATATGATGAATCGT	1256
RESULT 14			
LOCUS	A40621	541 bp	DNA linear PAT 05-MAR-1997
DEFINITION	Sequence 21 from Patent WO9425601.		
ACCESSION	A40621		
VERSION	A40621.1	GI:2296656	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 541)		
AUTHORS	Maertens,G. and Stuyver,L.		
TITLE	NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS THERAPEUTIC AND DIAGNOSTIC AGENTS		
JOURNAL	Patent: WO 9425601-A 21 10-NOV-1994;		
COMMENT	INNOGENETICS NV (BE) Other publication CA 2139100 941110 Other publication AU 6722294 941121 Other publication CN 1108030 950906 Other publication FI 946066 941223 Other publication NO 944967 941221 Other publication JP 7508423T 950921.		
FEATURES	Location/Qualifiers		
source	1..541		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
	/clone="BR36-9-20"		
CDS	2..541		
	/note="unnamed protein product; protein sequence is in conflict with the conceptual translation"		
	/codon_start=1		
	/protein_id="CAA02497.1"		
	/db_xref="GI:2296657"		
	/translation="VGAPVGGVARALAHGVRALEDGINFATGNLPGCCFSIFLLALFS CLIPHPASLEWRNTSGLYVLTNDCSNSSI VYEADVVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTATASIRSHVDLLVGAATMCSALYVGMCGAVFLVQGAFTFRPRHQ TVQTCNCSLYPGHLSGRMA"		
BASE COUNT	106 a	156 c	142 g 137 t
ORIGIN			
Query Match	85.6%; Score 539.4; DB 6; Length 541;		
Best Local Similarity	99.8%; Pred. No. 9.7e-140;		
Matches	540; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	63	CGTCGGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGGCATGGCGTACGGGCCCTTGA	122
Db	1	CGTCGGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGGCATGGCGTACGGGCCCTTGA	60
Qy	123	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTCTTCT	182
Db	61	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTCTTCT	120
Qy	183	CGTCTGTCTCTTCTGTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGTG	242

Db	121	TGCTCTGTTTCTCTTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCCTGG	180
Qy	243	CCTCTATGTCTTACCAACGACTGTTCCAATAGCAGTATTGTGTACGAGCCGATGACGCT	302
Db	181	CCTCTATGTCTTACCAACGACTGTTCCAATAGCAGTATTGTGTACGAGCCGATGACGCT	240
Qy	303	TATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGGACGGCAATATACATCCACGTGCTG	362
Db	241	TATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGGACGGCAATATACATCCACGTGCTG	300
Qy	363	GACCCAGTGCACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACACCCGCTTCGATAACG	422
Db	301	GACCCAGTGCACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACACCCGCTTCGATAACG	360
Qy	423	CAGTCATGTGGACCTATTAGTGGCGCGGCACGATGTGCTCTGCGCTCTACGTGGGTGA	482
Db	361	CAGTCATGTGGACCTATTAGTGGCGCGGCACGATGTGCTCTGCGCTCTACGTGGGTGA	420
Qy	483	CATGTGTGGGCTGTCTTCTCGTGGGACAAAGCCTTCAAGTTCAGACCTCTGTCGCCATCA	542
Db	421	CATGTGTGGGCTGTCTTCTCGTGGGACAAAGCCTTCAAGTTCAGACCTCTGTCGCCATCA	480
Qy	543	AACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCGCATCTTTTACGAGCATCGAATGGC	602
Db	481	AACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCGCATCTTTTACGAGCATCGAATGGC	540
Qy	603	T 603	
Db	541	T 541	
RESULT 15			
LOCUS	AX031599	541 bp	DNA linear PAT 20-SEP-2000
DEFINITION	Sequence 21 from Patent EP1004670.		
ACCESSION	AX031599		
VERSION	AX031599.1	GI:10278836	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1		
AUTHORS	Maertens,G. and Stuyver,L.		
TITLE	Sequences of hepatitis c virus genotypes and their use as therapeutic and diagnostic agents		
JOURNAL	Patent: EP 1004670-A 21 31-MAY-2000;		
FEATURES	Location/Qualifiers		
source	1..541		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
	/clone="BR36-9-20"		
	<2..>541		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC09675.1"		
	/db_xref="GI:10278837"		
	/translation="VGAPVGGVARALAHGVRALEDGINFATGNLPGCCFSIFLLALFS CLIPHPASLEWRNTSGLYVLTNDCSNSSI VYEADVVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTATASIRSHVDLLVGAATMCSALYVGMCGAVFLVQGAFTFRPRHQ TVQTCNCSLYPGHLSGRMA"		
CDS	<2..>541		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC09675.1"		
	/db_xref="GI:10278837"		
	/translation="VGAPVGGVARALAHGVRALEDGINFATGNLPGCCFSIFLLALFS CLIPHPASLEWRNTSGLYVLTNDCSNSSI VYEADVVILHTPGCIPCVQDGNSTCWTPTVPTVAVKYVGATTATASIRSHVDLLVGAATMCSALYVGMCGAVFLVQGAFTFRPRHQ TVQTCNCSLYPGHLSGRMA"		
BASE COUNT	106 a	156 c	142 g 137 t
ORIGIN			
Query Match	85.6%; Score 539.4; DB 6; Length 541;		
Best Local Similarity	99.8%; Pred. No. 9.7e-140;		
Matches	540; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	63	CGTCGGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGGCATGGCGTACGGGCCCTTGA	122
Db	1	CGTCGGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGGCATGGCGTACGGGCCCTTGA	60
Qy	123	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTCTTCT	182
Db	61	AGACGGGATAAATTCGCAACAGGGAATTTGCCGGTGTCTCTTTTCTATTTTCTCTTCT	120
Qy	183	CGTCTGTCTCTTCTGTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGTG	242

QY 123 AGACGGGATAAATTTCGCAACAGGGAATTTGCCGGTTCCTCCTTTCTATTTCCTTCT 182
Db |||||
QY 61 AGACGGGATAAATTTCGCAACAGGGAATTTGCCGGTTCCTCCTTTCTATTTCCTTCT 120
Db |||||
QY 183 CGCTCTGTCTCTCTTCTTAATTATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCG 242
Db |||||
QY 121 TGCTCTGTCTCTCTTCTTAATTATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCG 180
Db |||||
QY 243 CCTCTATGTCCTTACCAACGACTGTTCCAAATAGCAGTATTGTACGAGGCCGATGACGT 302
Db |||||
QY 181 CCTCTATGTCCTTACCAACGACTGTTCCAAATAGCAGTATTGTACGAGGCCGATGACGT 240
Db |||||
QY 303 TATTCTGCACACACCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 362
Db |||||
QY 241 TATTCTGCACACACCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 300
Db |||||
QY 363 GACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACCGCTTCGATACG 422
Db |||||
QY 301 GACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACCGCTTCGATACG 360
Db |||||
QY 423 CAGTCATGTGGACCTATTAGTGGGGCGGCCACGATGTGCTCTGCGCTCTACGTGGGTGA 482
Db |||||
QY 361 CAGTCATGTGGACCTATTAGTGGGGCGGCCACGATGTGCTCTGCGCTCTACGTGGGTGA 420
Db |||||
QY 483 CATGTGTGGGGCTGTCTTCTCTGTGGGACAAGGCTTCAGTTTCAGACCTCGTCGCCATCA 542
Db |||||
QY 421 CATGTGTGGGGCTGTCTTCTCTGTGGGACAAGGCTTCAGTTTCAGACCTCGTCGCCATCA 480
Db |||||
QY 543 AACGGTCCAGACCTGTAACTGTCTCGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 602
Db |||||
QY 481 AACGGTCCAGACCTGTAACTGTCTCGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 540
Db |||||
QY 603 T 603
Db |
QY 541 T 541
Db |

Search completed: December 20, 2003, 02:02:07
Job time : 2387.36 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 174.169 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-29
Perfect score: 630
Sequence: 1 ATGGGTAGTCATCGATAC.....TGATGATCACTGGTATATAG 630

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_13Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	630	17	AAT12965 HCV E1 construct H
2	630	100.0	630	24	AAQ78036 Hepatitis C virus
3	536.2	85.1	541	15	AAQ78033 Hepatitis C virus
4	533	84.6	541	15	AAQ78032 Hepatitis C virus
5	504.2	80.0	541	15	AAQ78029 Hepatitis C virus
6	503.2	79.9	540	15	AAQ78031 Hepatitis C virus
7	502.6	79.8	541	15	AAQ78030 Hepatitis C virus
8	499.4	79.3	541	15	AAQ78034 Hepatitis C virus

9	499.4	79.3	541	15	AAQ78035 Hepatitis C virus
10	497.8	79.0	541	15	AAQ78036 Hepatitis C virus
11	383.6	60.9	576	16	AAQ83883 Hepatitis C virus
12	383.6	60.9	576	17	AAT16597 Hepatitis C virus
13	382	60.6	576	16	AAQ83882 Hepatitis C virus
14	382	60.6	576	17	AAT16596 Hepatitis C virus
15	380.4	60.4	576	16	AAQ83881 Hepatitis C virus
16	380.4	60.4	576	17	AAT16595 Hepatitis C virus
17	374	59.4	576	16	AAQ83880 Hepatitis C virus
18	374	59.4	576	17	AAT16594 Hepatitis C virus
19	368.2	58.4	576	16	AAQ83879 Hepatitis C virus
20	368.2	58.4	576	17	AAT16593 Hepatitis C virus
21	344.4	54.7	1734	14	AAQ40435 Hepatitis C virus
22	342.8	54.4	1734	14	AAQ40434 Hepatitis C virus
23	341.2	54.2	1734	14	AAQ40436 Hepatitis C virus
24	339.6	53.9	1734	14	AAQ40425 Hepatitis C virus
25	336	53.3	1734	14	AAQ40437 Hepatitis C virus
26	336	53.3	1734	14	AAQ40438 Hepatitis C virus
27	335.6	53.3	630	17	AAT12966 HCV E1 construct H
28	335.6	53.3	630	24	AAQ48930 Hepatitis C virus
29	334.8	53.1	1734	14	AAQ40427 Hepatitis C virus
30	334.8	53.1	9405	14	AAQ40426 Full-length Hepati
31	332.8	52.8	1734	14	AAQ40429 Hepatitis C virus
32	332.8	52.8	3461	15	AAQ64068 Non-A, non-B hepat
33	332.8	52.8	3461	16	AAT30386 S'UTR/CORE/ENV/NS1
34	330	52.4	2116	12	AAQ12242 Encodes PT-NANBH V
35	329.8	52.3	633	17	AAT12706 HCV E1 construct H
36	329.8	52.3	633	24	AAQ48915 Hepatitis C virus
37	329.6	52.3	1734	14	AAQ40430 Hepatitis C virus
38	328	52.1	1734	14	AAQ40428 Hepatitis C virus
39	326.8	51.9	932	13	AAQ20923 C10-E12 DNA fragme
40	326.8	51.9	1037	15	AAQ58450 Hepatitis C virus
41	326.8	51.9	1562	19	AAV60672 Fragment #5 isolat
42	326.8	51.9	1953	25	AAQ55222 Plasmid pIDK82 DNA
43	326.8	51.9	2033	15	AAQ64913 Hepatitis C virus
44	326.8	51.9	2033	15	AAQ86788 Hepatitis C virus
45	326.8	51.9	2829	19	AAQ60673 Fragment #6 isolat

ALIGNMENTS

RESULT 1
AAT12965
ID AAT12965 standard; DNA; 630 BP.
XX
AC AAT12965;
XX
XX 24-SEP-1996 (first entry)
DT
DE HCV E1 construct HCCI62.
XX
XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.
XX
OS Hepatitis C virus.
XX
XX WO9604385-A2.
PN
XX
PD 15-FEB-1996.
XX
XX 31-JUL-1995; 95WO-EP03031.
XX
XX 29-JUL-1994; 94EP-0870132.
PR
XX
XX (INNO-) INNOGENETICS NV.
PA
XX
XX Bosman F, Buyse M, De Martynoff G, Maertens G;
PI
XX
XX WPI; 1996-129401/13.
DR
XX
XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

Qy

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

1 234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950515253545556575859606162636465666768697071727374757677787980818283848586878889909192939495969798991001011021031041051061071081091101111121131141151161171181191201211221231241251261271281291301311321331341351361371381391401411421431441451461471481491501511521531541551561571581591601611621631641651661671681691701711721731741751761771781791801811821831841851861871881891901911921931941951961971981992002012022032042052062072082092102112122132142152162172182192202212222232242252262272282292302312322332342352362372382392402412422432442452462472482492502512522532542552562572582592602612622632642652662672682692702712722732742752762772782792802812822832842852862872882892902912922932942952962972982993003013023033043053063073083093103113123133143153163173183193203213223233243253263273283293303313323333343353363373383393403413423433443453463473483493503513523533543553563573583593603613623633643653663673683693703713723733743753763773783793803813823833843853863873883893903913923933943953963973983994004014024034044054064074084094104114124134144154164174184194204214224234244254264274284294304314324334344354364374384394404414424434444454464474484494504514524534544554564574584594604614624634644654664674684694704714724734744754764774784794804814824834844854864874884894904914924934944954964974984995005015025035045055065075085095105115125135145155165175185195205215225235245255265275285295305315325335345355365375385395405415425435445455465475485495505515525535545555565575585595605615625635645655665675685695705715725735745755765775785795805815825835845855865875885895905915925935945955965975985996006016026036046056066076086096106116126136146156166176186196206216226236246256266276286296306316326336346356366376386396406416426436446456466476486496506516526536546556566576586596606616626636646656666676686696706716726736746756766776786796806816826836846856866876886896906916926936946956966976986997007017027037047057067077087097107117127137147157167177187197207217227237247257267277287297307317327337347357367377387397407417427437447457467477487497507517527537547557567577587597607617627637647657667677687697707717727737747757767777787797807817827837847857867877887897907917927937947957967977987998008018028038048058068078088098108118128138148158168178188198208218228238248258268278288298308318328338348358368378388398408418428438448458468478488498508518528538548558568578588598608618628638648658668678688698708718728738748758768778788798808818828838848858868878888898908918928938948958968978988999009019029039049059069079089099109119129139149159169179189199209219229239249259269279289299309319329339349359369379389399409419429439449459469479489499509519529539549559569579589599609619629639649659669679689699709719729739749759769779789799809819829839849859869879889899909919929939949959969979989991000100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011011102110311041105110611071108110911101111111211131114111511161117111811191120112111221123112411251126112711281129113011311132113311341135113611371138113911401141114211431144114511461147114811491150115111521153115411551156115711581159116011611162116311641165116611671168116911701171117211731174117511761177117811791180118111821183118411851186118711881189119011911192119311941195119611971198119912001201120212031204120512061207120812091210121112121213121412151216121712181219122012211222122312241225122612271228122912301231123212331234123512361237123812391240124112421243124412451246124712481249125012511252125312541255125612571258125912601261126212631264126512661267126812691270127112721273127412751276127712781279128012811282128312841285128612871288128912901291129212931294129512961297129812991300

DT	25-MAR-2003 (updated)	243	CCTCTATGTCCTTACCACGAGCTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGACGT	302
DT	01-AUG-1995 (first entry)	181	CTCTATGTCCTTACCACGAGCTGTTCCAAATAGCAGTATTGTGTACGAGGCCGATGACGT	240
XX				
DE	Hepatitis C virus Core/E1 region.	303	TATTTCTGCACACACCCGGGTGCAATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG	362
XX				
KW	Hepatitis C virus; HCV; primer; probe; detection; diagnosis;	241	TATTTCTGCACACACCCGGGTGCAATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG	300
KW	classification; immunisation; prophylaxis; serotyping; ss.			
XX				
OS	Hepatitis C virus type 3a.	363	GACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACCGCTTCGATACG	422
XX				
FT	Key	301	GACCCAGTGACACCTTACAGTGGCAGTCAAGTACGTCGGAGCAACCAACCGCTTCGATACG	360
FT	2..541			
FT	/tag= a			
FT	/product= Core/E1 polypeptide.			
XX				
XX	WO9425601-A2.	423	CAGTCATGTGACCTATTAGTGGCGCGGCCACGATGTGCTGTGCGCTTACGTGGGTGA	482
XX				
XX	10-NOV-1994.	361	CAGTCATGTGACCTATTAGTGGCGCGGCCACGATGTGCTGTGCGCTTACGTGGGTGA	420
XX				
XX	27-APR-1994; 94WO-EP01323.	483	CATGTGTGGGGTGTCTTCTCTCTGGGACAAAGCTTCACGTTCCAGACCTCGTCGCCATCA	542
XX				
XX	27-APR-1993; 93EP-0401099.	421	TATGTGTGGGGTGTCTTCTCTCTGGGACAAAGCTTCACGTTCCAGACCTCGTCGCCATCA	480
XX	05-AUG-1993; 93EP-0402019.			
XX	(INNO-) INNOGENETICS NV SA.	543	AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCCATCTTTTCAGGACATCGAATGGC	602
XX				
XX	Maertens G, Stuyver L;	481	AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCCATCTTTTCAGGACATCGAATGGC	540
XX	WPI; 1994-358277/44.			
XX	P-PSDB; AAR63280.			
XX				
XX	New polynucleotide sequences from hepatitis C virus - and related			
XX	vectors, polypeptide(s) and antibodies, useful for immunisation,			
XX	treatment, diagnosis and typing of HCV isolates			
XX	Claim 2; Page 109-110; 404pp; English.			
XX				
XX	Compositions comprising at least 5, and pref. 8 or more contiguous			
XX	nucleotides selected from an HCV type 3 genomic sequence, more			
XX	particularly (i) the region spanning positions 417-957 of the			
XX	Core/E1 region of HCV subtype 3a; (ii) the region spanning positions			
XX	4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning			
XX	positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the			
XX	region spanning positions 8023-8235 of the NS5 region of the BR36			
XX	subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic			
XX	sequence, may be used as primers to amplify nucleic acid from an			
XX	isolate belonging to a specific genotype, or as a probe for specific			
XX	detection/classification of nucleic acid. Polypeptides encoded by			
XX	the nucleotides in such compositions may be used for immunisation			
XX	against HCV, for the detection of antibodies directed against HCV			
XX	and for serotyping. This sequence corresponds to the Core/E1			
XX	region of HCV subtype 3a and is taken from a clone designated			
XX	BR36-9-13.			
XX	(Updated on 25-MAR-2003 to correct PN field.)			
XX				
XX	Sequence 541 BP; 107 A; 155 C; 142 G; 137 T; 0 other;			
XX				
XX	Query Match 84.6%; Score 533; DB 15; Length 541;			
XX	Best Local Similarity 99.1%; Pred. No. 5.9e-161;			
XX	Matches 536; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
QY	63 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTCGCGATGCGGTGAGGGCCCTTGA	122		
DB	1 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTCGCGATGCGGTGAGGGCCCTTGA	60		
QY	123 AGACGGGATAAATTTTCGACAGGAAATTTGCCGGTGTCTCTTTCTATTTTCTTCT	182		
DB	61 AGACGGGATAAATTTTCGACAGGAAATTTGCCGGTGTCTCTTTCTATTTTCTTCT	120		
QY	183 CGCTCTGTCTCTCTCTTAATTCATCCAGCAGTGTAGAGTGGCGGAATACGTCGTGG	242		
DB	121 TGTCTGTGTCTCTCTCTTAATTCATCCAGCAGTGTAGAGTGGCGGAATACGTCGTGG	180		

RESULT 5

ID	AAQ78029	AAQ78029 standard; cDNA; 541 BP.
XX		
AC	AAQ78029;	
XX		
DT	25-MAR-2003 (updated)	
DT	20-JUL-1995 (first entry)	
XX		
DE	Hepatitis C virus Core/E1 region.	
XX		
KW	Hepatitis C virus; HCV; primer; probe; detection; diagnosis;	
KW	classification; immunisation; prophylaxis; serotyping; ss.	
XX		
OS	Hepatitis C virus type 3a.	
XX		
EH	Key	Location/Qualifiers
FT	CDS	2..541
FT		/tag= a
FT		/product= Core/E1 polypeptide.
XX		
PN	WO9425601-A2.	
XX		
PD	10-NOV-1994.	
XX		
PF	27-APR-1994; 94WO-EP01323.	
XX		
PR	27-APR-1993; 93EP-0401099.	
PR	05-AUG-1993; 93EP-0402019.	
XX		
PA	(INNO-) INNOGENETICS NV SA.	
XX		
PI	Maertens G, Stuyver L;	
XX		
DR	WPI; 1994-358277/44.	
DR	P-PSDB; AAR63277.	
XX		
PT	New polynucleotide sequences from hepatitis C virus - and related	
PT	vectors, polypeptide(s) and antibodies, useful for immunisation,	
PT	treatment, diagnosis and typing of HCV isolates	
XX		
PS	Claim 2; Page 103-104; 404pp; English.	
XX		
CC	Compositions comprising at least 5, and pref. 8 or more contiguous	

CC nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by the nucleotides in such compositions may be used for immunisation against HCV, for the detection of antibodies directed against HCV and for serotyping. This sequence corresponds to the Core/E1 region of HCV subtype 3a and is taken from a clone designated HD10-2-5.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 541 BP; 104 A; 153 C; 145 G; 139 T; 0 other;

Query Match 80.0%; Score 504.2; DB 15; Length 541;

Best Local Similarity 95.7%; Pred. No. 1.1e-151;

Matches 518; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 63 CGTCGGCGTCCGCTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGAGGGCCCTTGA 122

DB 1 CGTCGGCGTCTCTGTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGAGGGCCCTTGA 60

QY 123 AGACGGGATAAATTCGCAACAGGGAATTTGCCCGGTCTCTCTTTCTATTTCTCTCT 182

DB 61 AGACGGGATAAATTCGCAACAGGGAATTTGCCCGGTCTCTCTTTCTATTTCTCTCT 120

QY 183 CGCTCTGTCTCTGTCTTAAATTCATCCAGCAGTCTAGTCTAGAGTGCGGGAATACGTCGG 242

DB 121 TGCTCTGTCTCTGTCTTAAATTCATCCAGCAGTCTAGTCTAGAGTGCGGGAATACGTCGG 180

QY 243 CCTCTATGCTTACCAACGAGTCTTCCAAATAGCAGTATTGTGTACGAGCGCGATGACGT 302

DB 181 CCTCTATGCTTACCAACGAGTCTTCCAAATAGCAGTATTGTGTATGAGGCGCGATGACGT 240

QY 303 TATTCTGCACACACCGGCTGCATACCTTGTCTCAGGACGCGCAATACATCCACGTGCTG 362

DB 241 TATTCTGCACACACCGGCTGTACTTGTCTCAGGACGCGTATACATCTCGGTGCTG 300

QY 363 GACCCAGTGACCTTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAGC 422

DB 301 GACCCAGTGACCTTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAGC 360

QY 423 CAGTCATGTGGACCTATTAGTGGGCGCGCCACGATGCTCTCGGCTCTACGTGGGTGA 482

DB 361 CAGGCAATGTAGACATGTTGGTGGGCGCGCCACGATGCTCTCTCTACGTGGGTGA 420

QY 483 CATGTGTGGGCTGTCTTCTCGTGGGCAAGCCTTTCAGCTTCAGACCTCGTCGCCATCA 542

DB 421 TATGTGTGGGCGGTCTTCTCTGTGGGCAAGCCTTTCAGCTTCAGACCTCGTCGCCATCA 480

QY 543 AACGCTCCAGACCTGTAACTGCTCGTGTACCCAGGCGCATTTTCAGGACATCGAATGCG 602

DB 481 AACGCTCCAGACCTGTAACTGCTCACTGTACCCAGGCGCATTTTCAGGACACCGAATGCG 540

QY 603 T 603

DB 541 T 541

RESULT 6

AAQ78031

ID AAQ78031 standard; cDNA; 540 BP.

XX AAQ78031;

XX AAQ78031;

DT 25-MAR-2003 (updated)

DT 21-JUL-1995 (first entry)

XX

DE Hepatitis C virus Core/E1 region.

XX

KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis; classification; immunisation; prophylaxis; serotyping; ss.

XX

OS Hepatitis C virus type 3a.

XX

FH Key Location/Qualifiers

FT CDS 2..541

FT /*tag= a

FT /product= Core/E1 polypeptide.

XX

PN WO9425601-A2.

XX

PD 10-NOV-1994.

XX

PF 27-APR-1994; 94WO-EP01323.

XX

PR 27-APR-1993; 93EP-0401099.

PR 05-AUG-1993; 93EP-0402019.

XX

XX (INNO-) INNOGENETICS NV SA.

XX

PI Maertens G, Stuyver L;

XX

XX WPI; 1994-358277/44.

DR P-PSDB; AAR63279.

XX

XX New polynucleotide sequences from hepatitis C virus - and related vectors, polypeptide(s) and antibodies, useful for immunisation, treatment, diagnosis and typing of HCV isolates

XX

XX Claim 2; Page 107-108; 404pp; English.

XX

CC Compositions comprising at least 5, and pref. 8 or more contiguous nucleotides selected from an HCV type 3 genomic sequence, more particularly (i) the region spanning positions 417-957 of the Core/E1 region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the region spanning positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to amplify nucleic acid from an isolate belonging to a specific genotype, or as a probe for specific detection/classification of nucleic acid. Polypeptides encoded by the nucleotides in such compositions may be used for immunisation against HCV, for the detection of antibodies directed against HCV and for serotyping. This sequence corresponds to the Core/E1 region of HCV subtype 3a and is taken from a clone designated HD10-2-21.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 540 BP; 105 A; 153 C; 144 G; 138 T; 0 other;

Query Match 79.9%; Score 503.2; DB 15; Length 540;

Best Local Similarity 95.7%; Pred. No. 2.3e-151;

Matches 517; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 64 GTCGGCGCTCCGCTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGAGGGCCCTTGA 123

DB 1 GTCGGCGCTCTCTGTAGGAGCGTCGCAAGAGCCCTTGGCATGCGTGAGGGCCCTTGA 60

QY 124 GACGGGATAAATTCGCAACAGGGAATTTGCCCGGTCTCTCTTTCTATTTCTCTCTC 183

DB 61 GACGGGATAAATTCGCAACAGGGAATTTGCCCGGTCTCTCTTTCTATTTCTCTCTCT 120

QY 184 GCTCTGTTCTCTGTTGTTAATTCATCCAGCAGTCTAGTCTAGAGTGGCGGAATACGTCTGC 243

DB 121 GCTCTGTTCTCTGTTGTTAATTCATCCAGCAGTCTAGTCTAGAGTGGCGGAATACGTCTGC 180

QY 244 CTCTATGCTCTTACCAACGACTGTTTCCAAATAGCAGTATTGTGTACAGGCGCGATGCGTT 303

DB 181 CTCTAGCTCTTACCAACGACTGTTTCCAAATAGCAGTATTGTGTATGAGCGCGATGCGTT 240

```
QY 304 ATTCTGCACACACCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTGG 363
DB 241 ATTCTGCACACACCGGCTGTGTACCTTGTGTTCAGGACGGTAATACATCTGCGTGTGG 300
QY 364 ACCCAGTGCACCTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAGC 423
DB 301 ACCCCAGTGCACCTACAGTGGCAGTCAAGTACGTCGAGCAACACCGCTTCGATAGC 360
QY 424 AGTCATGTGGACCTATTAGTGGGCGCGCCACGATGTCTCTGCGCTCTACGTGGTGAC 483
DB 361 AGCAATGTAGACATATTGTGGGCGCGCCAGATGTCTCTGCTCTCTACGTGGTGAT 420
QY 484 ATGTGTGGGGCTGTCTTCTCTCTGTGGGCAAGCCTTCAGTTTCAGACCTCTGTCGCATCA 543
DB 421 ATGTGTGGGGCGCTCTTCTCTCTGTGGGCAAGCCTTCAGTTTCAGACCTCTGTCGCATCA 480
QY 544 ACGGTCCAGACCTGTAACTGCTCGTGTACCCAGGCCATCTTTTCAGGACATCGAATGCT 603
DB 481 ACGGTCCAGACCTGTAACTGCTCACTGTATACCCAGGCCATCTTTTCAGGACACCGAATGCT 540

RESULT 7
AAQ78030
ID AAQ78030 standard; cDNA; 541 BP.
XX AC
XX AAQ78030;
XX
XX 25-MAR-2003 (updated)
DT 21-JUL-1995 (first entry)
XX
DE Hepatitis C virus Core/E1 region.
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
XX Hepatitis C virus type 3a.
XX
XX Key Location/Qualifiers
FH 2..541
CDS /*tag= a
FT /product= Core/E1 polypeptide.
FT
XX
XX W09425601-A2.
XX
XX 10-NOV-1994.
XX
XX 27-APR-1994; 94WO-EP01323.
XX
XX 27-APR-1993; 93EP-0401099.
XX 05-AUG-1993; 93EP-0402019.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
XX Maertens G, Stuyver L;
XX
XX WPI; 1994-358277/44.
XX P-PSDB; AAR63278.
XX
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates
XX
XX Claim 2; Page 105-106; 404pp; English.
XX
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a, (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
```

```
CC sequence, may be used as primers to amplify nucleic acid from an
CC isolate belonging to a specific genotype, or as a probe for specific
CC detection/classification of nucleic acid. Polypeptides encoded by
CC the nucleotides in such compositions may be used for immunisation
CC against HCV, for the detection of antibodies directed against HCV
CC and for serotyping. This sequence corresponds to the Core/E1
CC region of HCV subtype 3a and is taken from a clone designated
CC HD10-2-14.
XX (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 541 BP; 106 A; 154 C; 143 G; 138 T; 0 other;
Query Match 79.8%; Score 502.6; DB 15; Length 541;
Best Local Similarity 95.6%; Pred. No. 3.5e-151;
Matches 517; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 63 CGTCGGCGCTCCCGTAGGAGCGCTCGCAAGACCCCTTGGCATGGCGTAGGGCCCTTGA 122
DB 1 CGTCGGCGCTCCCGTAGGAGCGCTCGCAAGACCCCTTGGCATGGCGTAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTCCCAACAGGGAATTTGCCCGGTTGCTCTCTTTCTATTTCTTCT 182
DB 61 AGACGGGATAAATTTCCCAACAGGGAATTTGCCCGGTTGCTCTCTTTCTATTTCTTCT 120
QY 183 CGCTCTGTCTTCTTGTCTTAATTCACAGCAGTAGCTAGAGTGGCGGAATACGTCTGG 242
DB 121 TGCTCTGTCTTCTTGTCTTAATTCACAGCAGTAGCTAGAGTGGCGGAACACGTCTGG 180
QY 243 CCTCTATGCTCTTACCAACGACTCTTCCAAATAGCAGTATTGTGTACGAGGCCGATGAGCT 302
DB 181 CCTCTATGCTCTTACCAACGACTCTTCCAAATAGCAGTATTGTGTATGAGGCCGATGAGCT 240
QY 303 TATTCTGCACACACCCGCTGCATACCTTGTGTCCAGGACGGCAATACATCCAGTGTGTG 362
DB 241 TATTCTGCACACACCCGCTGTGTACCTTGTGTTCAGGACGGTAATACATCTGCGTGTG 300
QY 363 GACCCAGTGCACCTACAGTGGCAGTCAAGTACGTGCGAGCAACCCGCTTCGATAGC 422
DB 301 GACCCAGTGCACCTACAGTGGCAGTCAAGTACGTGCGAGCAACCCGCTTCGATAGC 360
QY 423 CAGTCATGTGGACCTATTAGTGGGCGCGCCACGATGTCTTGGCTCTACGTGGGTGA 482
DB 361 CAGGCATGTAGACATATTGGTGGGCGCGCCACATGTGCTCTGTCTCTACGTGGGTGA 420
QY 483 CATGTGTGGGCTGTCTTCTCTCGTGGGCAAGCCCTTACGTTACAGCTCTGTCGCCATCA 542
DB 421 TATGTGTGGGCGCTCTTCTCTCGTGGGCAAGCCCTTACGTTACAGCTCTGTCGCCATCA 480
QY 543 AACGGTCCAGACCTGTAACTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGCG 602
DB 481 AACGGTCCAGACCTGTAACTGCTGTACCCAGGCCATCTTTTCAGGACATCGAATGCG 540
QY 603 T 603
DB 541 T 541

RESULT 8
AAQ78034
ID AAQ78034 standard; cDNA; 541 BP.
XX AC
XX AAQ78034;
XX
XX 25-MAR-2003 (updated)
DT 01-AUG-1995 (first entry)
XX
XX Hepatitis C virus Core/E1 region.
DE
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX
XX Hepatitis C virus type 3a.
XX
```

FH Key Location/Qualifiers
 FT CDS 2..541
 FT /*tag= a
 FT /product= Core/E1 polypeptide.
 XX W09425601-A2.
 PN 10-NOV-1994.
 PD 27-APR-1994; 94WO-EP01323.
 XX 27-APR-1993; 93EP-0401099.
 XX 05-AUG-1993; 93EP-0402019.
 XX (INNO-) INNOGENETICS NV SA.
 XX Maertens G, Stuyver L;
 PI WPI; 1994-358277/44.
 DR P-PSDB; AAR63282.
 XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates
 XX Claim 2; Page 113-114; 404pp; English.
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the
 CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
 CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
 CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
 CC region spanning positions 8023-8235 of the NS5 region of the BR36
 CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
 CC sequence, may be used as primers to amplify nucleic acid from an
 CC isolate belonging to a specific genotype, or as a probe for specific
 CC detection/classification of nucleic acid. Polypeptides encoded by
 CC the nucleotides in such compositions may be used for immunisation
 CC against HCV, for the detection of antibodies directed against HCV
 CC and for serotyping. This sequence corresponds to the Core/E1
 CC region of HCV subtype 3a and is taken from a clone designated
 CC BR33-1-10.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 541 BP; 100 A; 157 C; 148 G; 136 T; 0 other;
 SQ Query Match 79.3%; Score 499.4; DB 15; Length 541;
 Best Local Similarity 95.2%; Pred. No. 3.8e-150;
 Matches 515; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 QY 63 CGTCGGCGCTCCCGTAGGAGCGGTGCGAAGAGCCCTTGGCGATGGCGTGAGGGCCCTTGA 122
 DB 1 CGTCGGCGCTCCCGTAGGAGCGGTGCGAAGAGCCCTTGGCGATGGCGTGAGGGCCCTTGA 60
 QY 123 AGACGGGATAAATTTGCGAACAGGGAATTTGCCGGTTGCTCTTTCTATTCTTCTCT 182
 DB 61 GGACGGGATAAATTTGCGAACAGGGAATTTGCCGGTTGCTCTTTCTATTCTTCTCT 120
 QY 183 CGCTCTGTTCTTCTGTTAATTCATCCAGCAGTAGCTAGAGTGCGCGAATACGTCG 242
 DB 121 TGCCTCTGTTCTTCTGTTAATTCATCCAGCAGTAGCTAGAGTGCGCGAATACGTCG 180
 QY 243 CCTCTATGTCCTTACCAGACGTGTTCCAAATAGCAGTATTGTACGAGCGCGATGACGT 302
 DB 181 CCTCTATGTCCTTACCAGACGTGTTCCAAATAGTATTGTATGAGCGCGATGACGT 240
 QY 303 TATTTCTGCACACACCGCGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGCTG 362
 DB 241 TATTTCTGCACGCGCGCGGTGTGTACCTTGTGTCCAGGACGGCAATACATGCTGCTG 300
 QY 363 GACCCGAGTGACCTTACAGTGCGAGTCAAGTACGTGCGAGCAACACCGCTTCGATACG 422

Db 301 GACCCAGTAACACCTACAGTGGCAGTACAGTACAGTACAGTACAGTACAGTACAGTACG 360
 QY 423 CAGTCATGTGGACCTATTAGTGGCGCGCGCCACGATGTCTCTGCGCTCTAGCTGGGTGA 482
 Db 361 CAGTCATGTGGACCTATTAGTGGCGCGCGCCACGATGTCTCTGCGCTTACGTGGGTGA 420
 QY 483 CATGTGTGGGCTGTCTTCTCTGTTGGGCAAGCCTTACGTTACAGCTTGTGCGCATCA 542
 Db 421 TATGTGTGGGCGGTCTTCTCTGTTGGGCAAGCCTTACGTTACAGCTTGTGCGCATCA 480
 QY 543 AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 602
 Db 481 AACGGTCCAGACCTGTAACTGCTGCTGTACCCAGGCCATCTTTCAGGACATCGAATGGC 602
 QY 603 T 603
 Db 541 T 541
 RESULT 9
 AAQ78035
 ID AAQ78035 standard; cDNA; 541 BP.
 XX AC AAQ78035;
 XX DT 25-MAR-2003 (updated)
 DT 01-AUG-1995 (first entry)
 DE Hepatitis C virus Core/E1 region.
 XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping; ss.
 XX OS Hepatitis C virus type 3a.
 XX PH Key Location/Qualifiers
 FT CDS 2..541
 FT /*tag= a
 FT /product= Core/E1 polypeptide.
 XX W09425601-A2.
 XX PD 10-NOV-1994.
 XX 27-APR-1994; 94WO-EP01323.
 XX 27-APR-1993; 93EP-0401099.
 XX 05-AUG-1993; 93EP-0402019.
 XX (INNO-) INNOGENETICS NV SA.
 XX Maertens G, Stuyver L;
 PI WPI; 1994-358277/44.
 DR P-PSDB; AAR63283.
 XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates
 XX Claim 2; Page 115-116; 404pp; English.
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the
 CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
 CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
 CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
 CC region spanning positions 8023-8235 of the NS5 region of the BR36
 CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
 CC sequence, may be used as primers to amplify nucleic acid from an
 CC isolate belonging to a specific genotype, or as a probe for specific
 CC detection/classification of nucleic acid. Polypeptides encoded by

Mon Dec 22 13:28:45 2003

us-09-899-303a-29.rng

/product= Core/E1 polypeptide.

CC the nucleotides in such compositions may be used for immunisation
CC against HCV, for the detection of antibodies directed against HCV
CC and for serotyping. This sequence corresponds to the Core/E1
CC region of HCV subtype 3a and is taken from a clone designated
CC BR33-1-19.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX

SQ Sequence 541 BP; 100 A; 155 C; 148 G; 138 T; 0 other;

Query Match 79.3%; Score 499.4; DB 15; Length 541;
Best Local Similarity 95.2%; Pred. No. 3.8e-150; Indels 0; Gaps 0;
Matches 515; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 63 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGCATGCGGTGAGGGCCCTTGA 122
DB 1 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGCATGCGGTGAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTCTCT 182
DB 61 GGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTCTCT 120
QY 183 CGCTCTGTCTCTTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGG 242
DB 121 TGTCTGTCTCTTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGG 180
QY 243 CCTCTATGTCCTTACCAAGAGCTGTTTCCAAATAGCAGTATTTGTACGAGCCGATGACGT 302
DB 181 CCTCTATGTCCTTACCAAGAGCTGTTTCCAAATAGTATTTGTATGAGCCGATGACGT 240
QY 303 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACGCAATACATCCAGTCGTG 362
DB 241 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACGCAATACATCCAGTCGTG 300
QY 363 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTCGAGGCAACACCCGCTTCGATACG 422
DB 301 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTCGAGGCAACACCCGCTTCGATACG 360
QY 423 CAGTCATGTGGAGCTTATAGTGGCGGCGGACGATGCTGCGCTTTACGTCGGTGA 482
DB 361 CAGTCATGTGGAGCTTATAGTGGCGGCGGACGATGCTGCGCTTTACGTCGGTGA 420
QY 483 CAGTCATGTGGAGCTTATAGTGGCGGCGGACGATGCTGCGCTTTACGTCGGTGA 542
DB 421 TATGTGTGGGCGGCTTCTCTCTGCGGACAAAGCTTCACGTTTCAGACCTCGTCGCCATCA 480
QY 543 AACGGTCCAGACCTGTAACTGCTGTGTATCCAGGCAATCTTTACGAGCAATCGAATGCG 602
DB 481 AACGGTCCAGACCTGTAACTGCTGTGTATCCAGGCAATCTTTACGAGCAATCGAATGCG 540
QY 603 T 603
DB 541 T 541

RESULT 10

AAQ78036
ID AAQ78036 standard; cDNA; 541 BP.
XX

AC AAQ78036;

XX 25-MAR-2003 (updated)
DT 01-AUG-1995 (first entry)

XX Hepatitis C virus Core/E1 region.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.
XX

OS Hepatitis C virus type 3a.

XX Key Location/Qualifiers

FT CDS 2..541

FT /*tag= a

FT
XX
PN W09425601-A2.
XX
PD 10-NOV-1994.
XX
PF 27-APR-1994; 94WO-EP01323.
XX
PR 27-APR-1993; 93EP-0401099.
PR 05-AUG-1993; 93EP-0402019.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI Maertens G, Stuyver L;
DR WPI; 1994-358277/44.
DR P-PSDB; AAR63284.
XX
PT New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates
XX
PS Claim 2; Page 117-118; 404pp; English.
CC
CC Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
CC sequence, may be used as primers to amplify nucleic acid from an
CC isolate belonging to a specific genotype, or as a probe for specific
CC detection/classification of nucleic acid. Polypeptides encoded by
CC the nucleotides in such compositions may be used for immunisation
CC against HCV, for the detection of antibodies directed against HCV
CC and for serotyping. This sequence corresponds to the Core/E1
CC region of HCV subtype 3a and is taken from a clone designated
CC BR33-1-20.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 541 BP; 100 A; 154 C; 148 G; 139 T; 0 other;

Query Match 79.0%; Score 497.8; DB 15; Length 541;
Best Local Similarity 95.0%; Pred. No. 1.2e-149;
Matches 514; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 63 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGCATGCGGTGAGGGCCCTTGA 122
DB 1 CGTCGGCGCTCCCGTAGGAGCGTCGCAAGAGCCCTTGCATGCGGTGAGGGCCCTTGA 60
QY 123 AGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTCTCT 182
DB 61 GGACGGGATAAATTTGCAACAGGGAATTTGCCGGTTCCTCTTTCTATTTCTCTCT 120
QY 183 CGCTCTGTCTCTTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGG 242
DB 121 TGTCTGTCTCTTCTTAAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAATACGTCGG 180
QY 243 CCTCTATGTCCTTACCAAGAGCTGTTTCCAAATAGCAGTATTTGTACGAGCCGATGACGT 302
DB 181 CCTCTATGTCCTTACCAAGAGCTGTTTCCAAATAGTATTTGTATGAGCCGATGACGT 240
QY 303 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACGCAATACATCCAGTCGTG 362
DB 241 TATTTCTGCACACACCCCGGCTGCATACCTTTGTTCAGGACGCAATACATCCAGTCGTG 300
QY 363 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTCGAGGCAACACCCGCTTCGATACG 422
DB 301 GACCCAGTGCACACCTACAGTGCAGTCAAGTACGTCGAGGCAACACCCGCTTCGATACG 360
QY 423 CAGTCATGTGGAGCTTATAGTGGCGGCGGACGATGCTGCTGCGCTCTACGTCGGTGA 482

```

Db 361 CAGTCATGTGGACCTGTTAGTAGGCGCGCCACGATGCTCTCGGCTTTTACGTGGTGA 420
Qy 483 CATGTGTGGGCTGTCTTCTCTGTGGGCAAGCCTTCACGTTTCAGACCTTCGTCGGCCATCA 542
Db 421 TATGTGTGGGCGCTTCTCTGTGGGCAAGCCTTCACGTTTCAGACCTTCGTCGGCCATCA 480
Qy 543 AACGCTCCAGACTGTAACTGCTCGCTGTACCGAGCCATCTTTCAGGACATCGAATGGC 602
Db 481 AACGCTCCAGACTGTAACTGCTCGCTGTACCGAGCCATCTTTCAGGACATCGAATGGC 540
Qy 603 T 603
Db 541 T 541

RESULT 11
AAQ83883
ID AAQ83883 standard; cDNA; 576 BP.
XX
AC AAQ83883;
XX
XX 25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX
DE Hepatitis C virus envelope 1 gene cDNA isolate S54.
XX
XX Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate S54;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..576
FT /*tag= a
XX
PN WO9501442-A2.
XX
PD 12-JAN-1995.
XX
PF 28-JUN-1994; 94WO-US07320.
XX
PR 29-JUN-1993; 93US-0086428.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
DR WPI; 1995-061006/08.
DR P-PSDB; AAR69672.
XX
XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX
PS Claim 1; Pages 72-73; 186pp; English.
XX
XX AAQ83883 encodes AAR69672 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate S54, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 576 BP; 110 A; 156 C; 163 G; 147 T; 0 other;

Query Match 60.9%; Score 383.6; DB 16; Length 576;
Best Local Similarity 96.6%; Pred. No. 6.9e-113;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 220 CTAGAGTGGCGGAATACGTCCTGGCTCTATGCTTACCAAGACTGTTCGAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGTCCTGGCTCTATGCTTACCAAGACTGTTCGAATAGCAGT 60

```

```

Qy 280 ATTGTGTACGAGCGCCGATGACGTTATTCTGCAACACCCGGCTGCATACCTTTGTGTCCAG 339
Db 61 ATTGTGTATGAGCGCGATGACGTCATTCTGCACACACCCGGCTGTGTACCTTTGTGTTCAG 120
Qy 340 GACGGCAATACATCCACGTCGTGGACCCCGAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCCACGTCGTGGACCCCGAGTGACACCTACGTCGTCAGGTACGTC 180
Qy 400 GGAGCAACCCACGCTTCGATACGACGTCATGTGGACCTATTAGTGGGCGCGCCACGATG 459
Db 181 GGAGCAACCCACGCTTCGATACGACGTCATGTGGACCTATTAGTGGGCGCGCCACGCTG 240
Qy 460 TGCTCTGGCTCTACGTGGGTGACATGTGTGGGCTGTCTTCTCTGTGGGCAAAAGCCTTC 519
Db 241 TGCTCTGGCTCTATGTGTGGGTGATATGTGTGGGCGCTCTTTCTCGTGGGCAAAAGCCTTC 300
Qy 520 ACGTTTCAGACCTTCGTCGCCATCAAAAGGTCAGACCTGTAACTGCTCGTGTACCCAGGC 579
Db 301 ACGTTTCAGACCTTCGTCGCCATCAAAAGGTCAGACCTGTAACTGCTCGTGTACCCAGGC 360
Qy 580 CATCTTTTCAGGACATCGAATGCTTTGGGATATGATGATGAATGCT 625
Db 361 CATCTTTTCAGGACATCGAATGCTTTGGGATATGATGATGAATGCT 406

RESULT 12
AAT16597
ID AAT16597 standard; cDNA; 576 BP.
XX
AC AAT16597;
XX
XX 30-SEP-1996 (first entry)
DT Hepatitis C virus isolate S54 envelope 1 gene.
DE
XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1..576
FT /*tag= a
FT /product= envelope-1 protein
FT /note= "does not contain start or stop codon"
XX
PN WO9605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US10398.
XX
PR 15-AUG-1994; 94US-0290665.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH ) US SEC DEPT HEALTH.
PA
XX PI Bukh J, Miller RH, Purcell RH;
XX
XX WPI; 1996-139709/14.
XX P-PSDB; AAR89543.
XX
PT DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX
XX Claim 1; Page 102; 340pp; English.
XX
XX AAT16559-TJ16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The

```

DR WPI; 1995-061006/08.
DR P-FSDB; AAR69671.

XX

DR P-P

```
XX 15-AUG-1995; 95WO-US10398.
PF 15-AUG-1994; 94US-0290665.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (USSH ) US SEC DEPT HEALTH.
PA PA
XX PI Bukh J, Miller RH, Purcell RH;
XX WPI; 1996-139709/14.
DR P-PSDB; AAR89542.
XX DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX Claim 1; Page 101-102; 340pp; English.
XX AAT16559-T16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The
CC proteins may also be used to detect antibodies against HCV in serum,
CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
CC used in the prevention of HCV infection.
XX Sequence 576 BP; 108 A; 153 C; 166 G; 149 T; 0 other;
XX Query Match 60.6%; Score 382; DB 17; Length 576;
XX Best Local Similarity 96.3%; Pred. No. 2.3e-112;
XX Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCACT 279
DB 1 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCACT 60
QY 280 ATTGTGTAGAGGCGGATGACGCTTATTCGACACACCCCGCTGCATACCTTGTGTCCAG 339
DB 61 ATTGTGTATGAGGCGGATGACGCTTATTCGACACACCCCGCTGTGTACCTTGTGTTCAG 120
QY 340 GACGCAATACATCCACGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
DB 121 GACGCAATACATCCACGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180
QY 400 GGAGCAACACCCGCTTCGATAGCGAGTCAATAGTGGGCGGCGCCAGCATG 459
DB 181 GGAGCAACACCCGCTTCGATAGCGAGTCAATAGTGGGCGGCGCCAGCATG 240
QY 460 TGCTCTGGGCTCTACGTGGGTGACATGTGTGGGCGTCTTCTCTGTGGGCAAGCCCTTC 519
DB 241 TGCTCTGGGCTCTACGTGGGTGACATGTGTGGGCGTCTTCTCTGTGGGCAAGCCCTTC 300
QY 520 ACGTTCAGACCTCGGCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 579
DB 301 ACGTTCAGACCTCGGCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 360
QY 580 CATCTTTGAGGACATCGAATGCTTGGGATATGATGATGAATGGT 625
DB 361 CATCTTTGAGGACATCGAATGCTTGGGATATGATGATGAATGGT 406
RESULT 15
AAQ83881
ID AAQ83881 standard; cDNA; 576 BP.
XX AAQ83881;
AC AAQ83881;
XX 25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX Hepatitis C virus envelope 1 gene cDNA isolate S2.
XX
```

```
KW Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate S2;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
XX Hepatitis C virus.
OS
XX Key Location/Qualifiers
FH mat_peptide 1..576
FT /*tag= a
XX
XX WO9501442-A2.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07320.
XX 29-JUN-1993; 93US-0086428.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Bukh J, Miller RH, Purcell RH;
XX WPI; 1995-061006/08.
XX P-PSDB; AAR69670.
XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX Claim 1; Page 71; 186pp; English.
XX AAQ83881 encodes AAR69670 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate S2, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 576 BP; 109 A; 156 C; 164 G; 147 T; 0 other;
XX Query Match 60.4%; Score 380.4; DB 16; Length 576;
XX Best Local Similarity 96.1%; Pred. No. 7.4e-112;
XX Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCACT 279
DB 1 CTAGAGTGGCGGAATACGCTCGGCCTCTATGCTTACCAACGACTGTTCCAAATAGCACT 60
QY 280 ATTGTGTAGAGGCGGATGACGCTTATTCGACACACCCCGCTGCATACCTTGTGTCCAG 339
DB 61 ATTGTGTATGAGGCGGATGACGCTTATTCGACACACCCCTGTGTACCTTGTGTTCAG 120
QY 340 GACGCAATACATCCACGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
DB 121 GACGCAATACATCCACGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTATGTC 180
QY 400 GGAGCAACACCCGCTTCGATAGCGAGTCAATAGTGGGCGGCGCCAGCATG 459
DB 181 GGAGCAACACCCGCTTCGATAGCGAGTCAATAGTGGGCGGCGCCAGCATG 240
QY 460 TGCTCTGGGCTCTACGTGGGTGACATGTGTGGGCGTCTTCTCTGTGGGCAAGCCCTTC 519
DB 241 TGCTCTGGGCTCTACGTGGGTGACATGTGTGGGCGTCTTCTCTGTGGGCAAGCCCTTC 300
QY 520 ACGTTCAGACCTCGGCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 579
DB 301 ACGTTCAGACCTCGGCCATCAACGGTCCAGACCTGTAACCTGCTCGTGTACCCAGGC 360
QY 580 CATCTTTGAGGACATCGAATGCTTGGGATATGATGATGAATGGT 625
DB 361 CATCTTTGAGGACATCGAATGCTTGGGATATGATGATGAATGGT 406
Search completed: December 19, 2003, 18:51:23
```

Job time : 176.169 secs

THIS PAGE BLANK (USP)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1620.71 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303A-29

Perfect score: 630

Sequence: 1 ATGGGTAAAGTCATCGATAC.....TGATGATGAACCTGGTAATAG 630

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estmu:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	53.8	8.5	488	9	AV755731
C 2	40.2	6.4	492	9	AV758366
C 3	40.2	6.4	633	28	AZ991479
C 4	38.4	6.1	521	10	BE756035

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 5	37.8	6.0	423	13	BW161485
C 6	37.8	6.0	701	13	BW152423
C 7	37.4	5.9	1252	12	BI517398
C 8	37	5.9	398	10	BE428299
C 9	36.8	5.8	356	13	BU862543
C 10	36.8	5.8	544	14	CA925476
C 11	36.8	5.8	568	14	CA925020
C 12	36.8	5.8	569	14	CA924833
C 13	36.8	5.8	577	14	CA924327
C 14	36.8	5.8	578	13	BU881720
C 15	36.6	5.8	465	13	BW153281
C 16	36.6	5.8	474	13	BW203948
C 17	36.6	5.8	491	13	BW156686
C 18	36.6	5.8	493	13	BW156832
C 19	36.6	5.8	497	13	BW282803
C 20	36.6	5.8	525	13	BW286366
C 21	36.6	5.8	550	13	BW162791
C 22	36.6	5.8	550	13	BW286224
C 23	36.6	5.8	551	13	BW158031
C 24	36.6	5.8	581	13	BW157153
C 25	36.6	5.8	582	13	BW162569
C 26	36.6	5.8	586	13	BW287549
C 27	36.6	5.8	618	13	BW157565
C 28	36.6	5.8	661	9	AV874950
C 29	36.6	5.8	661	13	BW152433
C 30	36.6	5.8	681	13	BW162273
C 31	36.6	5.8	687	13	BW150128
C 32	36.6	5.8	688	13	BW159151
C 33	36.6	5.8	690	13	BW154364
C 34	36.6	5.8	691	13	BW163099
C 35	36.6	5.8	698	9	AV865274
C 36	36.6	5.8	706	13	BW150559
C 37	36.6	5.8	706	13	BW162414
C 38	36.6	5.8	707	13	BW150580
C 39	36.6	5.8	709	13	BW159160
C 40	36.6	5.8	710	13	BW160242
C 41	36.6	5.8	711	13	BW152595
C 42	36.6	5.8	716	13	BW152623
C 43	36.6	5.8	716	13	BW152657
C 44	36.6	5.8	718	13	BW152737
C 45	36.6	5.8	718	13	BW162324

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
L., Xu,S., Gu,W., Tu,Y., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguaguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

Location/Qualifiers

BASE CO
ORIGIN

Mon Dec 22 13:28:47 2003

QY 352 TCCACGTGCTGGACCCAGTGCACCTACAGTGGCAGTCAAGTACGTTCGGAGCAACAC 411
 Db 311 ACNGANGTCAGAAAGCAGCAGTGTCTTACTACACCAACAGATGACGACACAG 252
 QY 412 GCTTCGATACGACGTATGTCGACCTATTAGTGGCGCGCCACGATGTGCTCTGGCGTC 471
 Db 251 ACTGCTATGACGCTTGGCCCATCTTAATATTCCTGGGTTCGCCAGGTACCCGCGAG 192
 QY 472 TACGTGGGTGACATGTGTGGGCTGTCTTCTCTGTGGGACAAGCCTTC 519
 Db 191 TCCGAGCTTTTCCAGGAGCGCAGTGTGTCCGAATCAGCAGACTCC 144

RESULT 6
 BW152423/c
 LOCUS
 DEFINITION
 BW152423 Nori Satoh unpublished cDNA library, gonad Ciona
 intestinalis cDNA clone rcig018h05 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.
 1 (bases 1 to 701)
 Satoh, Y., Shin-i, T., Kohara, Y. and Satoh, N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
 source
 1..701
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcig018h05"
 /tissue_type="gonad"
 /clone_lib="Nori Satoh unpublished cDNA library, gonad"
 182 a 166 c 185 g 165 t 3 others

BASE COUNT
 ORIGIN
 Query Match 6.0%; Score 37.8; DB 13; Length 701;
 Best Local Similarity 48.3%; Pred. No. 35;
 Matches 102; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 292 GCCGATGACGTTATTCGACACACCCGCGCTGCATACCTTGTTCGAGGACGCAATACA 351
 Db 370 GCAGACGGTGTATTCTGCNCTCGGCTTCCTCGGTACCCGATACCGATGACGTCAATAAC 311
 QY 352 TCCACGTGTGGACCCGACGTCACCTACAGTGGCAGTCAAGTACGTTCGGAGCAACAC 411
 Db 310 ACTGACGTCAGAAAGCAGCAGTGTCTTACTACAAACCAACGATGACGACACAG 251
 QY 412 GCTTCGATACGACGTATGTGGACCTATTAGTGGCGCGCCACGATGTGCTCTGGCGTC 471
 Db 250 ACTGCCATGATGACGGTGTGGCCCATCTTATATATTCCTGGGTTCGCCAGGTACCCGCGAT 191
 QY 472 TACGTGGGTGACATGTGTGGGCTGTCTTCC 502
 Db 190 TCCGAGCTTTTCCCTGGAGCGGCGAGTGTTC 160

RESULT 7
 BW157398/c
 LOCUS
 DEFINITION
 BW157398
 603041654F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5163013 5',
 mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Homo sapiens (human)
 Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 1252)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Prepared by: the I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11404 row: k column: 14
 High quality sequence stop: 524.
 Location/Qualifiers

FEATURES
 source
 1..1252
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5163013"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_116"
 /notes="Organ: pooled colon, kidney, stomach; Vector: pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."

BASE COUNT
 ORIGIN
 Query Match 5.9%; Score 37.4; DB 12; Length 1252;
 Best Local Similarity 49.7%; Pred. No. 48;
 Matches 95; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 314 CACCGGCTGCATACCTTGTTCGAGGACGCAATACATCCACGCTTGGACCCAGTGA 373
 Db 200 CACCACTCCAAATACCTTGACACGAGGAGACCGGTAGGAACTCGGATCCACAGGG 141
 QY 374 CACCTACAGTGGCAGTCAAGTACGTCGGAGCAACACCGCTTCGATACGAGTCATGTGG 433
 Db 140 CAGGTGAGGCAAGTCCACATCTTCGGAGCAAGCACCTTCACCTCCGGCTCGCGGA 81
 QY 434 ACCTATTATGTGGCGCGCCACGATGTGCTCTGGCTCTACGTGGGTGACATGTGTGGGG 493
 Db 80 GCCAGGCAATGTAGCGCGAGTCTCAGAGGACAGAGCTGGAGCCGCGACATGGCTGGAG 21
 QY 494 CTGTCTTCTTC 504
 Db 20 CGCTGGCGCCG 10

RESULT 8
 BE428299/c
 LOCUS
 DEFINITION
 BE428299
 MTD005.E02F90616 ITEC MTD Durum Wheat Root Library Triticum
 turgidum subsp. durum cDNA clone MTD005.E02, mRNA sequence.
 BE428299
 BE428299.1 GI:9426142
 EST.

BE428299
 MTD005.E02F90616 ITEC MTD Durum wheat linear EST 26-JUL-2000
 turgidum subsp. durum cDNA clone MTD005.E02, mRNA sequence.

Unneberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.
 The poplar tree transcriptome: Analysis of expressed sequence tags
 from multiple libraries
 Unpublished
 Contact: BHALERAO RUPALI R.
 Umea Plant Science Center
 Department of Plant Physiology
 University of Umea, 901 87 Umea, Sweden
 Tel: +46 90 786 5279
 Fax: +46 90 786 6676
 Email: rupali.bhalerao@plantphys.umu.se.
 Location/Qualifiers
 1. 336
 /organism="Populus tremula x Populus tremuloides"
 /mol_type="mRNA"
 /db_xref="taxon:47664"
 /tissue_type="imbibed seed"
 /clone_lib="Populus imbibed seed cDNA library"
 70 a 59 c 122 g 104 t 1 others
 BASE COUNT
 ORIGIN

Query Match 5.8%; Score 36.8; DB 13; Length 356;
 Best Local Similarity 50.6%; Pred.No.59;
 Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 169 TCTATTTTCCTTCGCTCTGTTCTCTCTGTTTAATTCATCAGCAGCTAGCTAGAGTGG 228
 Db 80 TGTATTATTCATTCAGTTTAAATTCGTTCTGTTATTCGTCACGAAGTCTAGGACGTG 139
 Qy 229 CGGAATACGTCGGCTCTATGTCCTTACCAAGACGTGTTCCAATACAGATATTGTGTAC 288
 Db 140 CCCAAAGACCTGGCCTCAATGATGACCAAGAACTTCATCGGTATGCGGTAGGTGGC 199
 Qy 289 GAGCGCATGACGTATTCTGCACACACCGCGTGCATACCTTGTGCCAGGACGG 344
 Db 200 TTGCTGGAGTTGGTGTCTGCCAACTTGGCGGTGTTCTGTCGCTCGGTGG 255

RESULT 10
 CA925476/c
 LOCUS
 DEFINITION MTU7TL.P1.F10 Aspen leaf cDNA Library Populus tremuloides cDNA,
 mRNA sequence.
 CA925476
 CA925476.1 GI:27412406
 EST.
 SOURCE Populus tremuloides (quaking aspen)
 ORGANISM Populus tremuloides
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 eurosid1; Malpighiales; Salicaceae; Populus.
 REFERENCE 1 (bases 1 to 544)
 AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
 ,C-J.
 Expressed sequence tags from Aspen
 Unpublished
 Contact: Tsai C-J
 Plant Biotech Research Center
 Michigan Technological University, School of Forest Resources &
 Environmental Science
 1400 Townsend Drive, Houghton, MI 49931-1295, USA
 Tel: 906 487 2914
 Fax: 906 487 2915
 Email: chtsai@mtu.edu.
 Location/Qualifiers
 1. 544
 /organism="Populus tremuloides"
 /mol_type="mRNA"
 /db_xref="taxon:3693"
 /clone_lib="Aspen leaf cDNA Library"
 /note="Organ: leaf"
 184 a 162 c 101 g 97 t
 BASE COUNT
 ORIGIN

```

LOCUS          CA924833          569 bp      mRNA      linear      EST 27-DEC-2002
DEFINITION     MTUTL.P10.B12 Aspen leaf cDNA Library Populus tremuloideis cDNA,
ACCESSION      CA924833
VERSION        CA924833.1      GI:27411763
KEYWORDS       EST.
SOURCE         Populus tremuloideis (quaking aspen)
ORGANISM       Populus tremuloideis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid1; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 569)
AUTHORS        Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
C-J.
EXPRESSED      Expressed sequence tags from Aspen
JOURNAL        Unpublished
COMMENT        Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES             source
     1..569
     /organism="Populus tremuloideis"
     /mol_type="mRNA"
     /db_xref="taxon:3693"
     /clone_lib="Aspen leaf cDNA Library"
     /note="Organ: leaf"
BASE COUNT        183 a 181 c 113 g 92 t
ORIGIN
Query Match       5.8%; Score 36.8; DB 14; Length 569;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 169 TCTATTTTCCTTCGCTCTGTTCTTCTTGTGTTAATTCATCCAGCAGTCTAGAGTGG 228
      |||||
Db 513 TGTATTGCATCTAGTTTAAATTTGCTCTTTGTTATTGCTCAAGCAAGTCTAGGACGTG 454

QY 229 CGGAATAGCTTGGCCTCTATGCTTCTTACCAACGCTTCCATACAGTATTGTGTAC 288
      |||||
Db 453 CCCAAGACGCTGGCCTCAATGACCAAGAACCTCATCGCATATGGCGTGTAGGTGGC 394

QY 289 GAGCGCATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTGCCAGGCGG 344
      |||||
Db 393 TTTGCTGGAGTTGGTGTCTGCCAACCTTGGCGGTGTGCTGGCTCGGTGG 338

RESULT 13
LOCUS          CA924327/c          577 bp      mRNA      linear      EST 27-DEC-2002
DEFINITION     MTUTL.P4.D04 Aspen leaf cDNA Library Populus tremuloideis cDNA,
ACCESSION      CA924327
VERSION        CA924327.1      GI:27411257
KEYWORDS       EST.
SOURCE         Populus tremuloideis (quaking aspen)
ORGANISM       Populus tremuloideis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid1; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 577)
AUTHORS        Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
C-J.
EXPRESSED      Expressed sequence tags from Aspen
JOURNAL        Unpublished
COMMENT        Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES             source
     1..568
     /organism="Populus tremuloideis"
     /mol_type="mRNA"
     /db_xref="taxon:3693"
     /clone_lib="Aspen leaf cDNA Library"
     /note="Organ: leaf"
BASE COUNT        183 a 181 c 113 g 91 t
ORIGIN
Query Match       5.8%; Score 36.8; DB 14; Length 568;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 169 TCTATTTTCCTTCGCTCTGTTCTTGTGTTAATTCATCCAGCAGTCTAGAGTGG 228
      |||||
Db 513 TGTATTGCATCTAGTTTAAATTTGCTCTTTGTTATTGCTCAAGCAAGTCTAGGACGTG 454

QY 229 CGGAATAGCTTGGCCTCTATGCTTCTTACCAACGCTTCCATACAGTATTGTGTAC 288
      |||||
Db 453 CCCAAGACGCTGGCCTCAATGACCAAGAACCTCATCGCATATGGCGTGTAGGTGGC 394

QY 289 GAGCGCATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTGCCAGGCGG 344
      |||||
Db 393 TTTGCTGGAGTTGGTGTCTGCCAACCTTGGCGGTGTGCTGGCTCGGTGG 338

RESULT 12
LOCUS          CA924833/c          569 bp      mRNA      linear      EST 27-DEC-2002
DEFINITION     MTUTL.P12.F05 Aspen leaf cDNA Library Populus tremuloideis cDNA,
ACCESSION      CA925020
VERSION        CA925020.1      GI:27411950
KEYWORDS       EST.
SOURCE         Populus tremuloideis (quaking aspen)
ORGANISM       Populus tremuloideis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid1; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 569)
AUTHORS        Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
C-J.
EXPRESSED      Expressed sequence tags from Aspen
JOURNAL        Unpublished
COMMENT        Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES             source
     1..568
     /organism="Populus tremuloideis"
     /mol_type="mRNA"
     /db_xref="taxon:3693"
     /clone_lib="Aspen leaf cDNA Library"
     /note="Organ: leaf"
BASE COUNT        183 a 181 c 113 g 91 t
ORIGIN
Query Match       5.8%; Score 36.8; DB 14; Length 568;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 169 TCTATTTTCCTTCGCTCTGTTCTTGTGTTAATTCATCCAGCAGTCTAGAGTGG 228
      |||||
Db 513 TGTATTGCATCTAGTTTAAATTTGCTCTTTGTTATTGCTCAAGCAAGTCTAGGACGTG 454

QY 229 CGGAATAGCTTGGCCTCTATGCTTCTTACCAACGCTTCCATACAGTATTGTGTAC 288
      |||||
Db 453 CCCAAGACGCTGGCCTCAATGACCAAGAACCTCATCGCATATGGCGTGTAGGTGGC 394

QY 289 GAGCGCATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTGCCAGGCGG 344
      |||||
Db 393 TTTGCTGGAGTTGGTGTCTGCCAACCTTGGCGGTGTGCTGGCTCGGTGG 338

RESULT 11
LOCUS          CA925020/c          568 bp      mRNA      linear      EST 27-DEC-2002
DEFINITION     MTUTL.P12.F05 Aspen leaf cDNA Library Populus tremuloideis cDNA,
ACCESSION      CA925020
VERSION        CA925020.1      GI:27411950
KEYWORDS       EST.
SOURCE         Populus tremuloideis (quaking aspen)
ORGANISM       Populus tremuloideis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid1; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 568)
AUTHORS        Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai
C-J.
EXPRESSED      Expressed sequence tags from Aspen
JOURNAL        Unpublished
COMMENT        Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES             source
     1..568
     /organism="Populus tremuloideis"
     /mol_type="mRNA"
     /db_xref="taxon:3693"
     /clone_lib="Aspen leaf cDNA Library"
     /note="Organ: leaf"
BASE COUNT        183 a 181 c 113 g 91 t
ORIGIN
Query Match       5.8%; Score 36.8; DB 14; Length 568;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 169 TCTATTTTCCTTCGCTCTGTTCTTGTGTTAATTCATCCAGCAGTCTAGAGTGG 228
      |||||
Db 513 TGTATTGCATCTAGTTTAAATTTGCTCTTTGTTATTGCTCAAGCAAGTCTAGGACGTG 454

QY 229 CGGAATAGCTTGGCCTCTATGCTTCTTACCAACGCTTCCATACAGTATTGTGTAC 288
      |||||
Db 453 CCCAAGACGCTGGCCTCAATGACCAAGAACCTCATCGCATATGGCGTGTAGGTGGC 394

QY 289 GAGCGCATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTGCCAGGCGG 344
      |||||
Db 393 TTTGCTGGAGTTGGTGTCTGCCAACCTTGGCGGTGTGCTGGCTCGGTGG 338

```

1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsa@mtu.edu.
Location/Qualifiers

FEATURES
source

1. 577
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen leaf cDNA Library"
/note="Organ: leaf"

BASE COUNT 192 a 175 c 112 g 98 t

ORIGIN

Query Match 5.8%; Score 36.8; DB 14; Length 577;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Qy 169 TCTATTTTCCTTCGCTCTGCTCTCTTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGG 228
Db 546 TGTATTGCATCTAGTTTAAATGCTCTTGTATTGCTCAAGCAAGTGTAGGACGTG 487
Qy 229 CGGAATACGCTGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTAC 288
Db 486 CCCAAGACGCTGCCCTCAATGACCAAGAACCTCATCGCATATGGCGGTGTAGTGGC 427
Qy 289 GAGCGCATGACGTTATTCGCACACACCCCGCTGCATACCTTGTGTCCAGGACGG 344
Db 426 TTTGCTGAGTGTGGTGTCTGCCAAACCTTGGCGGTGTGTGTGGCTCGGTGG 371

RESULT 14

LOCUS

BU881720 578 bp mRNA linear EST 16-OCT-2002
UM66TF02 Populus flower cDNA library Populus balsamifera subsp.

DEFINITION

BU881720 trichocarpa cDNA 5 prime, mRNA sequence.

ACCESSION

BU881720.1 GI:24073244

VERSION

KEYWORDS

SOURCE

ORGANISM

Populus balsamifera subsp. trichocarpa
Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; rosids
; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: BHALERAO RUPALI R.
Umea Plant Science Center
Department of Plant Physiology
University of Umea, 901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: rupali.bhalerao@plantphys.umu.se.
Location/Qualifiers

FEATURES

source

1. 578
/organism="Populus balsamifera subsp. trichocarpa"
/mol_type="mRNA"
/sub_species="trichocarpa"
/db_xref="taxon:3694"
/clone_lib="Populus flower cDNA library"
/note="Organ: flower"

BASE COUNT 121 a 107 c 153 g 197 t

ORIGIN

Query Match 5.8%; Score 36.8; DB 13; Length 578;
Best Local Similarity 50.6%; Pred. No. 62;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Qy 169 TCTATTTTCCTTCGCTCTGCTCTTCTTGTCTTAATTCATCCAGCAGCTAGTCTAGAGTGG 228

Db 81 TGTATTGCATCTAGTTTAAATGCTCTTGTATTGCTCAAGCAAGTGTAGGACGTG 140
Qy 229 CGGAATACGCTGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGTATTGTGTAC 288
Db 141 CCCAAGACGCTGGCTCAATGACCAAGAACTTTCATCGCATATGGCGGTGTAGTGGC 200
Qy 289 GAGCGCATGACGTTATTCGCACACACCCCGCTGCATACCTTGTGTCCAGGACGG 344
Db 201 TTTGCTGAGTGTGGTGTCTGCCAAACCTTGGCGGTGTGTGTGGCTCGGTGG 256

RESULT 15

LOCUS

BM153281/1 465 bp mRNA linear EST 03-NOV-2002
BM153281 Nori Satoh unpublished cDNA library, gonad Ciona
intestinalis cDNA clone rcigd020b10 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BM153281.1 GI:24510506
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 465)
Satou, F., Shin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished
Contact: Nori Satoh
Department of Zoology
Kyoto University
Kyoto 606-8502, Japan
Sakyo-ku, Kyoto, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers

FEATURES

source

1. 465
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcigd020b10"
/tissue type="gonad"
/clone_lib="Nori Satoh unpublished cDNA library, gonad"

BASE COUNT 121 a 109 c 123 g 111 t 1 others

Query Match 5.8%; Score 36.6; DB 13; Length 465;
Best Local Similarity 48.3%; Pred. No. 68;
Matches 102; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 292 GCCGATGACGTTATTCTGCACACACCCGGGTGCATACCTTGTCTCCAGGACGCAATACA 351

Db 373 GCAGACGCTGTTATTCTCGCTCGCTTCTCGGTACCCGATACCGATGACGTCATAC 314

Qy 352 TCCAGGTGCTGGACCCCGACCTACAGTGGGAGTCAAGTACGTCGGAGCAACACC 411

Db 313 ACTGACGTCAGAAAGCGACGATCAGCTGTACTACACCAACACAGATGACGCAACACG 254

Qy 412 GCTTCGATACGACGTCATGTGGACCTATTAGTGGCGGCGGACGATGCTCTGCGCTC 471

Db 253 ACTGCCATAGTGAGCGTTGGGCGCCCATCTTTATATTCGCGGTTCCAGGATACCCCGCAT 194

Qy 472 TACGTGGGTGACATGTGTGGGGCTGTCTTCC 502

Db 193 TCCGAGCTTTTCCCTGGAGCGGAGTGTTC 163

Search completed: December 20, 2003, 06:54:55
Job time : 1624.71 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.4354 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-29
Perfect score: 630
Sequence: 1 ATGGTAGGTCATGCATAC.....TGATGATCAACTGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	630	US-08-612-973-29	Sequence 29, Appl
2	630	100.0	630	US-08-927-597-29	Sequence 29, Appl
3	383.6	60.9	576	US-08-086-428B-39	Sequence 39, Appl
4	383.6	60.9	576	US-08-468-570-39	Sequence 39, Appl
5	383.6	60.9	576	US-08-290-665A-39	Sequence 39, Appl
6	383.6	60.9	576	US-08-466-601A-39	Sequence 39, Appl
7	383.6	60.9	576	PCT-US95-10398-39	Sequence 39, Appl
8	382	60.6	576	US-08-086-428B-38	Sequence 38, Appl
9	382	60.6	576	US-08-468-570-38	Sequence 38, Appl
10	382	60.6	576	US-08-290-665A-38	Sequence 38, Appl
11	382	60.6	576	US-08-466-601A-38	Sequence 38, Appl
12	382	60.6	576	PCT-US95-10398-38	Sequence 38, Appl
13	380.4	60.4	576	US-08-086-428B-37	Sequence 37, Appl
14	380.4	60.4	576	US-08-468-570-37	Sequence 37, Appl
15	380.4	60.4	576	US-08-290-665A-37	Sequence 37, Appl
16	380.4	60.4	576	US-08-466-601A-37	Sequence 37, Appl
17	380.4	60.4	576	PCT-US95-10398-37	Sequence 37, Appl
18	374	59.4	576	US-08-086-428B-36	Sequence 36, Appl
19	374	59.4	576	US-08-468-570-36	Sequence 36, Appl
20	374	59.4	576	US-08-290-665A-36	Sequence 36, Appl
21	374	59.4	576	US-08-466-601A-36	Sequence 36, Appl
22	374	59.4	576	PCT-US95-10398-36	Sequence 36, Appl
23	368.2	58.4	576	US-08-086-428B-35	Sequence 35, Appl
24	368.2	58.4	576	US-08-468-570-35	Sequence 35, Appl
25	368.2	58.4	576	US-08-290-665A-35	Sequence 35, Appl
26	368.2	58.4	576	US-08-466-601A-35	Sequence 35, Appl
27	368.2	58.4	576	PCT-US95-10398-35	Sequence 35, Appl

28	335.6	53.3	630	3	US-08-612-973-31	Sequence 31, Appl
29	335.6	53.3	630	3	US-08-927-597-31	Sequence 31, Appl
30	330	52.4	2116	3	US-08-191-160-21	Sequence 21, Appl
31	329.8	52.3	633	3	US-08-612-973-7	Sequence 7, Appl
32	329.8	52.3	633	3	US-08-927-597-7	Sequence 7, Appl
33	326.8	51.9	932	1	US-08-081-072-15	Sequence 15, Appl
34	326.8	51.9	932	1	US-08-449-093A-15	Sequence 15, Appl
35	326.8	51.9	1037	1	US-08-462-195-3	Sequence 3, Appl
36	326.8	51.9	1037	2	US-08-636-883-3	Sequence 3, Appl
37	326.8	51.9	1037	3	US-09-127-829-3	Sequence 3, Appl
38	325.2	51.6	9595	3	US-09-014-416-4	Sequence 4, Appl
39	325.2	51.6	9599	3	US-09-014-416-6	Sequence 6, Appl
40	324.8	51.6	1167	1	US-08-324-977-9	Sequence 9, Appl
41	324.8	51.6	1167	2	US-08-384-616-9	Sequence 9, Appl
42	324.8	51.6	1167	2	US-08-904-686A-9	Sequence 9, Appl
43	324.8	51.6	1167	3	US-09-315-850-9	Sequence 9, Appl
44	324.8	51.6	1499	1	US-08-324-977-3	Sequence 3, Appl
45	324.8	51.6	1499	2	US-08-384-616-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-29
; Sequence 29, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..627
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624

QY 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGCTACCCAGGCCATCTTTTCAGGACATCGAATG 600
Db 541 CAAACGGTCCAGACCTGTAACTGCTCGCTGCTACCCAGGCCATCTTTTCAGGACATCGAATG 600
QY 601 GCTTGGGATATGATGATGAAGTGGTAATAG 630
Db 601 GCTTGGGATATGATGATGAAGTGGTAATAG 630
RESULT 3
US-08-086-428B-39
; Sequence 39, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S54
US-08-086-428B-39
Query Match 60.9%; Score 383.6; DB 1; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 220 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTCTTACCAACGACCTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTGCGCTCTATGCTCTTACCAACGACCTGTTCCAAATAGCAGT 60
QY 280 ATTGTGTACAGGCCGATACGCTTATCTGCACACACCGGCTGCATACCTGTGTCCAG 339
Db 61 ATTGTGTACAGGCCGATACGCTTATCTGCACACACCGGCTGTGTACCTGTGTTCAG 120
QY 340 GACGGCAATACATCCACGCTGCGCCAGCCAGTACACCTACAGTGGCGAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCCACGCTGCGCCAGCCAGTACACCTACAGTGGCGAGTCAAGTACGTC 180

QY 400 GGAGCAACACCGCTTCGATAGCGAGTCATGTGGACCTATTATAGTGGCGCGGCCAGCATG 459
Db 181 GGAGCAACACCGCTTCGATAGCGAGTCATGTGGACCTATTATAGTGGCGCGGCCAGCATG 240
QY 460 TGCCTCGCGCTCTACGTGGGTGACATGTGTGGGGCTGTCTTCTCGTGGGACAAAGCCTTC 519
Db 241 TGCCTCGCGCTCTATGTGGGTGATGTGTGGGGCGCTCTTCTCGTGGGACAAAGCCTTC 300
QY 520 ACGTTACAGACCTCGTGGCCATCAACGGTCCAGACCTGTAACCTGCTGTAACCCAGGC 579
Db 301 ACGTTACAGACCTCGTGGCCATCAACGGTCCAGACCTGTAACCTGCTGTAACCCAGGC 360
QY 580 CATCTTCAGGACATCGAATGCTTGGGATATGATGATGATGATGATGATGATGATGATGAT 625
Db 361 CATCTTCAGGACATCGAATGCTTGGGATATGATGATGATGATGATGATGATGATGATGAT 406
RESULT 4
US-08-468-570-39
; Sequence 39, Application US/08468570
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,570
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S54
US-08-468-570-39
Query Match 60.9%; Score 383.6; DB 2; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;


```
/ REFERENCE/DOCKET NUMBER: 2026-4070US2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ TELEX: 421792
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 576 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ ORIGINAL SOURCE:
/ ORGANISM: homosapiens
/ INDIVIDUAL ISOLATE: S54
/ US-08-466-601A-39

Query Match          60.9%; Score 383.6; DB 4; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 60

QY 280 ATTGTGTACGAGCCGATGACGCTTATCTGCACACACCCGGCTGCATACCTTGTGTCAG 339
Db 61 ATTGTGTATGAGCGCGATGACGCTTATCTGCACACACCCGGCTGTGTACCTTGTGTCAG 120

QY 340 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180

QY 400 GGAGCAACACCGCTTCGATACGCTGATGCGACCTTATTAGTGGCGCGCCAGCATG 459
Db 181 GGAGCAACACCGCTTCGATACGCTGATGCGACCTTATTAGTGGCGCGCCAGCATG 240

QY 460 TGCTCTGCGCTCTACGCTGGGTGACATGTTGGGGGTGCTCTCTCTGCGGACAAAGCCTTC 519
Db 241 TGCTCTGCGCTCTACGCTGGGTGATATGTTGGGGCGCTCTTCTCTGCGGACAAAGCCTTC 300

QY 520 ACGTTTCAGACCTCGTCGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 579
Db 301 ACGTTTCAGACCTCGTCGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 360

QY 580 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
Db 361 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406

RESULT 7
PCT-US95-10398-39
/ Sequence 39, Application PC/TUS9510398
/ GENERAL INFORMATION:
/ APPLICANT: BUKH, J., MILLER, R. H. AND
/ APPLICANT: PURCELL, R. H.
/ TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
/ TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
/ TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
/ TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
/ NUMBER OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
/ NUMBER OF SEQUENCES: 263
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORGAN & FINNEGAN
/ STREET: 345 PARK AVENUE
/ CITY: NEW YORK
/ STATE: NEW YORK
/ COUNTRY: USA
/ ZIP: 10154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WORDPERFECT 5.1
```

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/10398
/ FILING DATE: 15-AUG-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/086,428
/ FILING DATE: 29 JUNE 1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/290/665
/ FILING DATE: 15 AUGUST 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: RICHARD W. BORK
/ REGISTRATION NUMBER: 36,459
/ REFERENCE/DOCKET NUMBER: 2026-4116
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ TELEX: 421792
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 576 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ ORIGINAL SOURCE:
/ ORGANISM: homosapiens
/ INDIVIDUAL ISOLATE: S54
/ PCT-US95-10398-39

Query Match          60.9%; Score 383.6; DB 5; Length 576;
Best Local Similarity 96.6%; Pred. No. 2.9e-118;
Matches 392; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACTGTTCCAAATAGCAGT 60

QY 280 ATTGTGTACGAGCCGATGACGCTTATCTGCACACACCCGGCTGCATACCTTGTGTCAG 339
Db 61 ATTGTGTATGAGCGCGATGACGCTTATCTGCACACACCCGGCTGTGTACCTTGTGTCAG 120

QY 340 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC 180

QY 400 GGAGCAACACCGCTTCGATACGCTGATGCGACCTTATTAGTGGCGCGCCAGCATG 459
Db 181 GGAGCAACACCGCTTCGATACGCTGATGCGACCTTATTAGTGGCGCGCCAGCATG 240

QY 460 TGCTCTGCGCTCTACGCTGGGTGACATGTTGGGGGTGCTCTCTCTGCGGACAAAGCCTTC 519
Db 241 TGCTCTGCGCTCTACGCTGGGTGATATGTTGGGGCGCTCTTCTCTGCGGACAAAGCCTTC 300

QY 520 ACGTTTCAGACCTCGTCGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 579
Db 301 ACGTTTCAGACCTCGTCGCCATCAACCGTCCAGACCTGTAACCTGCTGCTGACCCAGGC 360

QY 580 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 625
Db 361 CATCTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGCT 406

RESULT 8
US-08-086-428B-38
/ Sequence 38, Application US/08086428B
/ Patent No. 5514539
/ GENERAL INFORMATION:
/ APPLICANT: BUKH, J., MILLER, R. H. AND
/ APPLICANT: PURCELL, R. H.
/ TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
/ TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
/ TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
/ TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
```

;; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES

;; NUMBER OF SEQUENCES: 159
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & FINNEGAN
;; STREET: 345 PARK AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/086,428B
;; FILING DATE: 29-JUN-1993

;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: RICHARD W. BORK
;; REGISTRATION NUMBER: 36,459
;; REFERENCE/DOCKET NUMBER: 2026-4070
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 758-4800
;; TELEFAX: (212) 751-6849
;; TELEX: 421792

;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 576 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; ORIGINAL SOURCE:
;; ORGANISM: homosapiens
;; INDIVIDUAL ISOLATE: S52
;; US-08-086-428B-38

Query Match 60.6%; Score 382; DB 1; Length 576;
Best Local Similarity 96.3%; Pred. No. 1e-117;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY	220	CTAGAGTGGCGGAAATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT	279
DB	1	CTAGAGTGGCGGAAATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT	60
QY	280	ATTGTGTACGAGCGGATGACGCTTATTCGACACACCCGGCTGCATACCTTGTGTCAG	339
DB	61	ATTGTGTATGAGCGGATGACGCTTATTCGACACACCCGGCTGTGTACCTTGTGTCAG	120
QY	340	GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC	399
DB	121	GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC	180
QY	400	GGAGCAACACCGCTTCGATACCGACTCATGTGGACCTATTAGTGGGCGGCGCCACGATG	459
DB	181	GGAGCAACACCGCTTCGATACCGACTCATGTGGACCTATTAGTGGGCGGCGCCACGATG	240
QY	460	TGCTCTGGGCTCTACGTTGGGTGACATGTGTGGGGTGTCCTTCCTCGTGGGCAAGCCTTC	519
DB	241	TGCTCTGGGCTCTATGTGGGTGATATGTGTGGGGCGGTCTTTCTCGTGGGCAAGCCTTC	300
QY	520	ACGTTTCAGACCTCGTGGCATCAACAGGTCAGACCTGTAACCTGCTGCTGACCCAGGC	579
DB	301	ACGTTTCAGACCTCGTGGCATCAACAGGTCAGACCTGTAACCTGCTGCTGACCCAGGC	360
QY	580	CATCTTTACGACATCGAATGGCTTGGGATATGATGATGAATGGT	625
DB	361	CATGTTTCAGACATCGAATGGCTTGGGATATGATGATGAATGGT	406

RESULT 9
US-08-468-570-38
; Sequence 38, Application US/08468570

;; Patent No. 5871962
;; GENERAL INFORMATION:
;; APPLICANT: BURKH, J., MILLER, R.H. AND
;; APPLICANT: PURCELL, R.H.
;; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
;; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
;; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
;; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
;; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
;; NUMBER OF SEQUENCES: 159
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & FINNEGAN
;; STREET: 345 PARK AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/468,570
;; FILING DATE: 6-JUN-1995

;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/086,428
;; FILING DATE: 29-JUN-1993
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: RICHARD W. BORK
;; REGISTRATION NUMBER: 36,459
;; REFERENCE/DOCKET NUMBER: 2026-4070US1

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 758-4800
;; TELEFAX: (212) 751-6849
;; TELEX: 421792
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 576 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; ORIGINAL SOURCE:
;; ORGANISM: homosapiens
;; INDIVIDUAL ISOLATE: S52
;; US-08-468-570-38

Query Match 60.6%; Score 382; DB 2; Length 576;
Best Local Similarity 96.3%; Pred. No. 1e-117;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY	220	CTAGAGTGGCGGAAATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT	279
DB	1	CTAGAGTGGCGGAAATACGTCCTGCTATGCTCTTACCAACGACTGTTCCCAATAGCAGT	60
QY	280	ATTGTGTACGAGCGGATGACGCTTATTCGACACACCCGGCTGCATACCTTGTGTCAG	339
DB	61	ATTGTGTATGAGCGGATGACGCTTATTCGACACACCCGGCTGTGTACCTTGTGTCAG	120
QY	340	GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC	399
DB	121	GACGGCAATACATCCAGCTGCTGGACCCAGTGACACCTACAGTGGCAGTCAAGTACGTC	180
QY	400	GGAGCAACACCGCTTCGATACCGACTCATGTGGACCTATTAGTGGGCGGCGCCACGATG	459
DB	181	GGAGCAACACCGCTTCGATACCGACTCATGTGGACCTATTAGTGGGCGGCGCCACGATG	240
QY	460	TGCTCTGGGCTCTACGTTGGGTGACATGTGTGGGGTGTCCTTCCTCGTGGGCAAGCCTTC	519
DB	241	TGCTCTGGGCTCTATGTGGGTGATATGTGTGGGGCGGTCTTTCTCGTGGGCAAGCCTTC	300

QY 520 ACCTTCAGACCTCGTCCGATCAACGGTCCAGACCTGTAACCTGCTCGCTGTACCCAGGC 579
Db 301 ACCTTCAGACCTCGTCCGATCAACGGTCCAGACCTGTAACCTGCTCGCTGTACCCAGGC 360
QY 580 CATCTTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGGT 625
Db 361 CATGTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGGT 406

RESULT 10

US-08-290-665A-38
; Sequence 38, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S52
US-08-290-665A-38

Query Match 60.6%; Score 382; DB 2; Length 576;
Best Local Similarity 96.3%; Pred. No. 1e-117;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGAATACGCTGCGCTCTATGCTTACCAAGACTGTTCCAAATAGCAGT 279
Db 1 CTAGAGTGGCGAATACGCTGCGCTCTATGCTTACCAAGACTGTTCCAAATAGCAGT 60
QY 280 ATTGTGTACGAGCCGATACGCTTATCTGCAACACCCGGGTGCATCCTGTGTCCAG 339
Db 61 ATTGTGTATGAGCCGATACGCTTATCTGCAACACCCGGGTGTGTACCTGTGTGTTTCA 120
QY 340 GACGGCAATACATCCAGCTGCTGGCCAGTGCACCTACAGTGCAGTCAAGTACGTC 399
Db 121 GACGGCAATACATCCAGCTGCTGGCCAGTGCACCTACAGTGCAGTCAAGTACGTC 180

QY 400 GGAGCAACCAACCGCTTCGATACGACCTGATGGAACCTATTATAGTGGCGCGCCACGATG 459
Db 181 GGAGCAACCAACCGCTTCGATACGACCTGATGGAACCTATTATAGTGGCGCGCCACGCTG 240
QY 460 TCCTCTGGGCTCTAGTGGGTGACATGTTGGGGCTGCTTCTCGTGGGCAAGCCTTC 519
Db 241 TCCTCTGGGCTCTAGTGGGTGATGTTGGGGCGCTTTCTCTGTTGGGCAAGCCTTC 300
QY 520 ACCTTCAGACCTCGTCCGATCAACGGTCCAGACCTGTAACCTGCTGTACCCAGGC 579
Db 301 ACCTTCAGACCTCGTCCGATCAACGGTCCAGACCTGTAACCTGCTGTACCCAGGC 360
QY 580 CATCTTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGGT 625
Db 361 CATGTTTCAGGACATCGAATGCTTGGGATATGATGATGAATGGT 406

RESULT 11

US-08-466-601A-38
; Sequence 38, Application US/08466601A
; Patent No. 6572864
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,601A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S52
US-08-466-601A-38

Query Match 60.6%; Score 382; DB 4; Length 576;
Best Local Similarity 96.3%; Pred. No. 1e-117;
Matches 391; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:

```



```

; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S2
US-08-086-428B-37

Query Match          60.4%; Score 380.4; DB 1; Length 576;
Best Local Similarity 96.1%; Pred. No. 3.4e-117;
Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACCTGTTCCAAATAGCAGT 279
DB 1 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACCTGTTCCAAATAGCAGT 60

QY 280 ATTGTGTACAGGCCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTTGTGTCAG 339
DB 61 ATTGTGTATGAGGCCGATGACGTTATTCTGCACACACCTGGCTGTGTACCTTTGTGTCAG 120

QY 340 GACGGCAATACATCCAGCTGTCGACCCCAAGTACACCTACAGTGGCAGTCAAGTACGTC 399
DB 121 GACGGTAATACATCCAGCTGTCGACCCCAAGTACACCTACAGTGGCAGTCAAGTATGTC 180

QY 400 GGAGCAACCAACCGCTTCGATAGCAGTCACTGACCTATTAGTGGCGCGGCACGATG 459
DB 181 GGAGCAACCAACCGCTTCGATAGCAGTCACTGACCTATTAGTGGCGCGGCACGATG 240

QY 460 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGCTGTCTTCTCTGTGGGCAAGCCTTC 519
DB 241 TGCTCTGCGCTCTACGTGGGTGATATGTTGGGGCGCTCTTCTCTGTGGGCAAGCCTTC 300

QY 520 ACGTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTCGCTACCCAGGC 579
DB 301 ACGTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTCGCTACCCAGGC 360

QY 580 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTT 625
DB 361 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTT 406

```

RESULT 14

```

US-08-468-570-37
; Sequence 37, Application US/08468570
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,570
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: S2
US-08-468-570-37

Query Match          60.4%; Score 380.4; DB 2; Length 576;
Best Local Similarity 96.1%; Pred. No. 3.4e-117;
Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 220 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACCTGTTCCAAATAGCAGT 279
DB 1 CTAGAGTGGCGGAATACGCTCGGCTCTATGCTTACCAACGACCTGTTCCAAATAGCAGT 60

QY 280 ATTGTGTACAGGCCGATGACGTTATTCTGCACACACCCGGCTGCATACCTTTGTGTCAG 339
DB 61 ATTGTGTATGAGGCCGATGACGTTATTCTGCACACACCTGGCTGTGTACCTTTGTGTCAG 120

QY 340 GACGGCAATACATCCAGCTGTCGACCCCAAGTACACCTACAGTGGCAGTCAAGTACGTC 399
DB 121 GACGGTAATACATCCAGCTGTCGACCCCAAGTACACCTACAGTGGCAGTCAAGTATGTC 180

QY 400 GGAGCAACCAACCGCTTCGATAGCAGTCACTGACCTATTAGTGGCGCGGCACGATG 459
DB 181 GGAGCAACCAACCGCTTCGATAGCAGTCACTGACCTATTAGTGGCGCGGCACGATG 240

QY 460 TGCTCTGCGCTCTACGTGGGTGACATGTTGGGGCTGTCTTCTCTGTGGGCAAGCCTTC 519
DB 241 TGCTCTGCGCTCTACGTGGGTGATATGTTGGGGCGCTCTTCTCTGTGGGCAAGCCTTC 300

QY 520 ACGTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTCGCTACCCAGGC 579
DB 301 ACGTTCAGACCTCGTCCCATCAACGGTCCAGACCTGTAACCTGCTCGCTACCCAGGC 360

QY 580 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTT 625
DB 361 CATCTTTAGGACATCGAATGCTTGGGATATGATGATGAATGTT 406

```

RESULT 15

```

US-08-290-665A-37
; Sequence 37, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

```

Mon Dec 22 13:28:46 2003

TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES

NUMBER OF SEQUENCES: 263
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,665A
 FILING DATE: 15-AUG-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: RICHARD W. BORK
 REGISTRATION NUMBER: 36,459
 REFERENCE/DOCKET NUMBER: 2026-4116
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 576 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: hominids
 INDIVIDUAL ISOLATE: S2
 US-08-290-665A-37

Query Match 60.4%; Score 380.4; DB 2; Length 576;
 Best Local Similarity 96.1%; Pred. No. 3.4e-117;
 Matches 390; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 220 CTAGAGTGGCGGATAGCTCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 279
 Db 1 CTAGAGTGGCGGATAGCTCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 60
 QY 280 ATTGTGTACGAGGCGGATAGCTCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 339
 Db 61 ATTGTGTATGAGGCGGATAGCTCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 120
 QY 340 GACGGCAATATACCTACGCTGCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 399
 Db 121 GACGGTAATATACCTACGCTGCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 180
 QY 400 GGAGCAACACCGCTTCGATACGCGATCATGTGACCTATTAGTGGGCGGCGCCACGATG 459
 Db 181 GGAGCAACACCGCTTCGATACGCGATCATGTGACCTATTAGTGGGCGGCGCCACGATG 240
 QY 460 TGCTCTGGGCTCTACGCTGCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 519
 Db 241 TGCTCTGGGCTCTACGCTGCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 300
 QY 520 ACGTTTCAGACCTCTGCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 579
 Db 301 ACGTTTCAGACCTCTGCTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 360
 QY 580 CATCTTTTCAGGACATGAGTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 625
 Db 361 CATCTTTTCAGGACATGAGTGGGCTCTATGTCCTTACCAACGACGTGTTCCAAATAGCAGT 406

Search completed: December 20, 2003, 07:03:16
 Job time : 46.4354 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:55:48 ; Search time 2385.36 Seconds
(without alignments)
10804.703 Million cell updates/sec

Title: US-09-899-303A-31
Perfect score: 630
Sequence: 1 ATGGGTAAAGTCATCGATAC.....TGATGATGAACGGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	630	100.0	630	6	A48693	A48693 Sequence 31
2	630	100.0	630	6	AR157341	AR157341 Sequence
3	630	100.0	630	6	AX452780	AX452780 Sequence
4	630	100.0	630	6	AX685032	AX685032 Sequence
5	602	95.6	957	14	HPCBE95A	L29577 Hepatitis C
6	602	95.6	959	6	A40649	A40649 Sequence 49
7	602	95.6	959	6	AX031627	AX031627 Sequence
8	602	95.6	959	6	AX031897	AX031897 Sequence
9	602	95.6	959	6	AX032167	AX032167 Sequence
10	602	95.6	959	6	BD172148	BD172148 New seque
11	600.8	95.4	959	6	A40653	A40653 Sequence 53
12	600.8	95.4	959	6	AX031631	AX031631 Sequence
13	600.8	95.4	959	6	AX031901	AX031901 Sequence
14	600.8	95.4	959	6	AX032171	AX032171 Sequence
15	600.8	95.4	959	6	BD172150	BD172150 New seque
16	597.2	94.8	959	6	A40651	A40651 Sequence 51
17	597.2	94.8	959	6	AX031629	AX031629 Sequence
18	597.2	94.8	959	6	AX031899	AX031899 Sequence
19	597.2	94.8	959	6	AX032169	AX032169 Sequence
20	597.2	94.8	959	6	BD172149	BD172149 New seque
21	580	92.1	580	6	A40645	A40645 Sequence 45
22	580	92.1	580	6	A40647	A40647 Sequence 47
23	580	92.1	580	6	AX031623	AX031623 Sequence
24	580	92.1	580	6	AX031625	AX031625 Sequence
25	580	92.1	580	6	AX031893	AX031893 Sequence
26	580	92.1	580	6	AX031895	AX031895 Sequence
27	580	92.1	580	6	AX032163	AX032163 Sequence
28	580	92.1	580	6	AX032185	AX032185 Sequence
29	580	92.1	580	6	BD172146	BD172146 New seque
30	580	92.1	580	6	BD172147	BD172147 New seque
31	573.2	91.0	579	6	A40753	A40753 Sequence 15
32	573.2	91.0	579	6	AX031731	AX031731 Sequence
33	573.2	91.0	579	6	AX032001	AX032001 Sequence
34	573.2	91.0	579	6	AX032271	AX032271 Sequence
35	573.2	91.0	579	6	BD172193	BD172193 New seque
36	573	91.0	578	14	HPCBE95F	L29582 Hepatitis C
37	539.6	85.7	1848	14	D50466	D50466 Hepatitis C
38	536	85.1	9408	14	AF064490	AF064490 Hepatitis C
39	524.8	83.3	9343	14	HCY1480	Y13184 Hepatitis C
40	510.2	81.0	579	6	A40755	A40755 Sequence 15
41	510.2	81.0	579	6	A40795	A40795 Sequence 19
42	510.2	81.0	579	6	AX031733	AX031733 Sequence
43	510.2	81.0	579	6	AX031773	AX031773 Sequence
44	510.2	81.0	579	6	AX032003	AX032003 Sequence
45	510.2	81.0	579	6	AX032043	AX032043 Sequence

ALIGNMENTS

RESULT 1
A48693
LOCUS A48693
DEFINITION Sequence 31 from Patent WO9604385.
ACCESSION A48693
VERSION A48693.1 GI:2302406
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 630)
AUTHORS Maertens,G., Bosman,F., De,M.G. and Buysse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 31 15-FEB-1996;

LOCUS	ARI57341	630 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 31 from patent US 6245503.				
ACCESSION	ARI57341				
VERSION	ARI57341.1				
KEYWORDS	GI:16218275				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 630)				
AUTHORS	Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.				
TITLE	Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use				
JOURNAL	Patent: US 6245503-A 31 12-JUN-2001;				
FEATURES	Location/Qualifiers				
source	1..630				
BASE COUNT	126 a	171 c	170 g	163 t	
ORIGIN					
Query Match	100.0%; Score 630; DB 6; Length 630;				
Best Local Similarity	100.0%; Pred. No. 1.9e-170;				
Matches	630; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
Qy	1	ATGGGTAAGGTCATCGATACCCCTAACCGTGGGATTCGCCGATCTCATGGGGTATATCCCG	60		
Db	1	ATGGGTAAGGTCATCGATACCCCTAACCGTGGGATTCGCCGATCTCATGGGGTATATCCCG	60		
Qy	61	CTCGTAGGCGGCCCATTTGGGGCGTTCGAAGGGCTCTCGCACACGGTGTGAGGGTCTT	120		
Db	61	CTCGTAGGCGGCCCATTTGGGGCGTTCGAAGGGCTCTCGCACACGGTGTGAGGGTCTT	120		
Qy	121	GAGGACGGGGTAAACTATGCAACAGGGAAATTTACCGGTTGCTCTTCTCTATCTTTATT	180		
Db	121	GAGGACGGGGTAAACTATGCAACAGGGAAATTTACCGGTTGCTCTTCTCTATCTTTATT	180		
Qy	181	CTTGCTCTTCTCTCGTGTCTGACCGTTCGGGCTCTGAGTTCCTACCGAAATGCCCTCT	240		
Db	181	CTTGCTCTTCTCTCGTGTCTGACCGTTCGGGCTCTGAGTTCCTACCGAAATGCCCTCT	240		
Qy	241	GGGATTTATCATGTTACCAATGATTTGCCAACTCTTCCATAGTCTATGAGCAGATAAC	300		
Db	241	GGGATTTATCATGTTACCAATGATTTGCCAACTCTTCCATAGTCTATGAGCAGATAAC	300		
Qy	301	CTGATCCTACACGACCTGGTTCGCTGCTGTCGTCATGACAGGTAATGTAGTAGATGC	360		
Db	301	CTGATCCTACACGACCTGGTTCGCTGCTGTCGTCATGACAGGTAATGTAGTAGATGC	360		
Qy	361	TGGGTCCAAATTTACCCCTACCTGTGACGAGCCCGAGCTCGGAGCAGTCAACGGCTCTCTT	420		
Db	361	TGGGTCCAAATTTACCCCTACCTGTGACGAGCCCGAGCTCGGAGCAGTCAACGGCTCTCTT	420		
Qy	421	CGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCTTATACGTAGGA	480		
Db	421	CGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCTTATACGTAGGA	480		
Qy	481	GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTCGCCAGCAC	540		
Db	481	GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTCGCCAGCAC	540		
Qy	541	GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG	600		
Db	541	GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG	600		
Qy	601	GCATGGGATATGATGATGAACTGGTAATAG	630		
Db	601	GCATGGGATATGATGATGAACTGGTAATAG	630		
RESULT 3					
LOCUS	AX452780				
DEFINITION	Sequence 31 from Patent EPI2111315.				
ACCESSION	AX452780				

COMMENT	INNOGENETICS NV (BE)	2172273	960215		
Other publication	CA 3382495	960304.			
Other publication	US 6245503				
Location/Qualifiers	1..630				
source	/organism="unidentified"				
CDS	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
	1..627				
BASE COUNT	126 a	171 c	163 t		
	ORIGIN				
Query Match 100.0%; Score 630; DB 6; Length 630;					
Best Local Similarity 100.0%; Pred. No. 1.9e-170;					
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	ATGGGTAAGGTCATCGATACCCCTAACCGTGGGATTCGCCGATCTCATGGGGTATATCCCG	60		
Db	1	ATGGGTAAGGTCATCGATACCCCTAACCGTGGGATTCGCCGATCTCATGGGGTATATCCCG	60		
Qy	61	CTCGTAGGCGGCCCATTTGGGGCGTTCGAAGGGCTCTCGCACACGGTGTGAGGGTCTT	120		
Db	61	CTCGTAGGCGGCCCATTTGGGGCGTTCGAAGGGCTCTCGCACACGGTGTGAGGGTCTT	120		
Qy	121	GAGGACGGGGTAAACTATGCAACAGGGAAATTTACCGGTTGCTCTTCTCTATCTTTATT	180		
Db	121	GAGGACGGGGTAAACTATGCAACAGGGAAATTTACCGGTTGCTCTTCTCTATCTTTATT	180		
Qy	181	CTTGCTCTTCTCTCGTGTCTGACCGTTCGGGCTCTGAGTTCCTACCGAAATGCCCTCT	240		
Db	181	CTTGCTCTTCTCTCGTGTCTGACCGTTCGGGCTCTGAGTTCCTACCGAAATGCCCTCT	240		
Qy	241	GGGATTTATCATGTTACCAATGATTTGCCAACTCTTCCATAGTCTATGAGCAGATAAC	300		
Db	241	GGGATTTATCATGTTACCAATGATTTGCCAACTCTTCCATAGTCTATGAGCAGATAAC	300		
Qy	301	CTGATCCTACACGACCTGGTTCGCTGCTGTCGTCATGACAGGTAATGTAGTAGATGC	360		
Db	301	CTGATCCTACACGACCTGGTTCGCTGCTGTCGTCATGACAGGTAATGTAGTAGATGC	360		
Qy	361	TGGGTCCAAATTTACCCCTACCTGTGACGAGCCCGAGCTCGGAGCAGTCAACGGCTCTCTT	420		
Db	361	TGGGTCCAAATTTACCCCTACCTGTGACGAGCCCGAGCTCGGAGCAGTCAACGGCTCTCTT	420		
Qy	421	CGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCTTATACGTAGGA	480		
Db	421	CGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCTTATACGTAGGA	480		
Qy	481	GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTCGCCAGCAC	540		
Db	481	GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTCGCCAGCAC	540		
Qy	541	GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG	600		
Db	541	GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG	600		
Qy	601	GCATGGGATATGATGATGAACTGGTAATAG	630		
Db	601	GCATGGGATATGATGATGAACTGGTAATAG	630		
RESULT 2					
LOCUS	ARI57341				

```
VERSION      AX452780.1  GI:21712465
KEYWORDS     Hepatitis C virus
SOURCE       Hepatitis C virus
ORGANISM     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
TITLE        Recombinant vectors for producing hcv envelope proteins
JOURNAL      Patent: EP 1211315-A 31 05-JUN-2002;
              Innogenetics N.V. (BE)
FEATURES     Location/Qualifiers
              1..630
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
              1..627
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD38077.1"
              /db_xref="GI:21712466"
              /translation="MGKVIDTLTCGFADLMGYIPLVGGPIGVVARALAHGVRLVDGV
              NYATGNLPGCSFISIFLALLSCLITVPASAVPYRNASGIYHVNDPCNSIVVEADNLI
              LHAPGCVPMWTGNVSRVQITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALYVG
              DACGALFLVGQMTYRPROHATVQNCNCSIYSHVTHGRMNDMMNNW"
              mat_peptide 1..624
              /product="unnamed"
BASE COUNT   126 a 171 c 170 g 163 t
ORIGIN
Query Match 100.0%; Score 630; DB 6; Length 630;
Best Local Similarity 100.0%; Pred. No. 1.9e-170;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTAAGGTATCGATACCCCTAAGTCGGATTCCGCGATTCGCGATCTCATGGGTATATCCCG 60
DB 1 ATGGGTAAGGTATCGATACCCCTAAGTCGGATTCCGCGATTCGCGATCTCATGGGTATATCCCG 60
QY 61 CTCGTAGGCGGCCCAATTCGGGCGCTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCT 120
DB 61 CTCGTAGGCGGCCCAATTCGGGCGCTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCT 120
QY 121 GAGGACGGGTAAGTAATGATGCAACAGGAAATTTACCGGTGCTCTTCTATCTTTATT 180
DB 121 GAGGACGGGTAAGTAATGATGCAACAGGAAATTTACCGGTGCTCTTCTATCTTTATT 180
QY 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGGCTCTGCGAGTTCCTTACCGAAATGCTCT 240
DB 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGGCTCTGCGAGTTCCTTACCGAAATGCTCT 240
QY 241 GGGATTTATCATGTTACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
DB 241 GGGATTTATCATGTTACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
QY 301 CTGATCCTACAGCACCTGGTTCGTCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
DB 301 CTGATCCTACAGCACCTGGTTCGTCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
QY 361 TGGGTCCAAATTTACCCCTACCTGTGTCAGCCCGAGCTTCGGAGCTCCCTACCGAATGCTCT 420
DB 361 TGGGTCCAAATTTACCCCTACCTGTGTCAGCCCGAGCTTCGGAGCTCCCTACCGAATGCTCT 420
QY 421 CGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCCGTTATACGTAGGA 480
DB 421 CGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCCGTTATACGTAGGA 480
QY 481 GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTTCGCCAGCAC 540
DB 481 GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTTCGCCAGCAC 540
QY 541 GCTACGGTGCAGAACTGCAACTGTTTCAATTTACAGTGGCCATGTTACCGGCCACCGGATG 600
DB 541 GCTACGGTGCAGAACTGCAACTGTTTCAATTTACAGTGGCCATGTTACCGGCCACCGGATG 600

RESULT 4
AX685032
LOCUS        AX685032
DEFINITION  Sequence 31 from Patent WO0205548.
ACCESSION   AX685032
VERSION     AX685032.1  GI:29371437
KEYWORDS    Hepatitis C virus
SOURCE      Hepatitis C virus
ORGANISM    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
              Hepacivirus.
REFERENCE    1
AUTHORS      Maertens,G., Bosman,F. and Buyse,M.A.
TITLE        Purified Hepatitis C Virus envelope proteins for diagnostic and
              therapeutic use
JOURNAL      Patent: WO 0205548-A 31 18-JUL-2002;
              INNOGENETICS N.V. (BE)
FEATURES     Location/Qualifiers
              1..630
              /organism="Hepatitis C virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:11103"
              1..627
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD86530.1"
              /db_xref="GI:29371438"
              /translation="MGKVIDTLTCGFADLMGYIPLVGGPIGVVARALAHGVRLVDGV
              NYATGNLPGCSFISIFLALLSCLITVPASAVPYRNASGIYHVNDPCNSIVVEADNLI
              LHAPGCVPMWTGNVSRVQITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALYVG
              DACGALFLVGQMTYRPROHATVQNCNCSIYSHVTHGRMNDMMNNW"
              mat_peptide 1..624
              /product="unnamed"
BASE COUNT   126 a 171 c 170 g 163 t
ORIGIN
Query Match 100.0%; Score 630; DB 6; Length 630;
Best Local Similarity 100.0%; Pred. No. 1.9e-170;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTAAGGTATCGATACCCCTAAGTCGGATTCCGCGATTCGCGATCTCATGGGTATATCCCG 60
DB 1 ATGGGTAAGGTATCGATACCCCTAAGTCGGATTCCGCGATTCGCGATCTCATGGGTATATCCCG 60
QY 61 CTCGTAGGCGGCCCAATTCGGGCGCTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCT 120
DB 61 CTCGTAGGCGGCCCAATTCGGGCGCTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCT 120
QY 121 GAGGACGGGTAAGTAATGATGCAACAGGAAATTTACCGGTGCTCTTCTATCTTTATT 180
DB 121 GAGGACGGGTAAGTAATGATGCAACAGGAAATTTACCGGTGCTCTTCTATCTTTATT 180
QY 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGGCTCTGCGAGTTCCTTACCGAAATGCTCT 240
DB 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGGCTCTGCGAGTTCCTTACCGAAATGCTCT 240
QY 241 GGGATTTATCATGTTACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
DB 241 GGGATTTATCATGTTACCAATGATGCCCAAACTCTTCCATAGTCTATGAGGCAGATAAC 300
QY 301 CTGATCCTACAGCACCTGGTTCGTCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
DB 301 CTGATCCTACAGCACCTGGTTCGTCGCTTGTGTGTCATGACAGGTAATGTAGTAGATGC 360
QY 361 TGGGTCCAAATTTACCCCTACCTGTGTCAGCCCGAGCTTCGGAGCTCCCTACCGAATGCTCT 420
DB 361 TGGGTCCAAATTTACCCCTACCTGTGTCAGCCCGAGCTTCGGAGCTCCCTACCGAATGCTCT 420
QY 421 CGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCCGTTATACGTAGGA 480
DB 421 CGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCGCTCTGCTCCCGTTATACGTAGGA 480
QY 481 GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTTCGCCAGCAC 540
DB 481 GACGCGTGTGGGCACTATTCTTGGTAGGCAAAATGTTCACTATAGGCTTCGCCAGCAC 540
QY 541 GCTACGGTGCAGAACTGCAACTGTTTCAATTTACAGTGGCCATGTTACCGGCCACCGGATG 600
DB 541 GCTACGGTGCAGAACTGCAACTGTTTCAATTTACAGTGGCCATGTTACCGGCCACCGGATG 600
```

QY	421	CGGAGAGCCGTTGACTACCTAGCGGAGGGCTGCCCTCTGCTCCGCGTTATACGTAGGA	480
Db	421	CGGAGAGCCGTTGACTACCTAGCGGAGGGCTGCCCTCTGCTCCGCGTTATACGTAGGA	480
QY	481	GACGCGTGTGGGGCACTATTCTTGTGAGGCCAAATGTTCACTATAGGCTCGCCAGCAC	540
Db	481	GACGCGTGTGGGGCACTATTCTTGTGAGGCCAAATGTTCACTATAGGCTCGCCAGCAC	540
QY	541	GCTACCGTGCAGAACTGCACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATG	600
Db	541	GCTACCGTGCAGAACTGCACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATG	600
QY	601	GCAATGGATATGATGAACTGCTAATAG	630
Db	601	GCAATGGATATGATGAACTGCTAATAG	630
RESULT 5			
LOCUS	HPCBE95A	957 bp ss-RNA	linear VRL 04-MAY-1995
DEFINITION	Hepatitis C virus (BE95) core protein gene, complete cds; envelope protein (E1) gene, complete cds.		
ACCESSION	L29577		
VERSION	L29577.1	GI:476661	
KEYWORDS	core protein; envelope protein.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1	(bases 1 to 957)	
AUTHORS	Stuyver, L., van Arnhem, W., Wyseur, A. and Maertens, G.		
TITLE	Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the hepatitis C virus type 5a		
JOURNAL	Biochem. Biophys. Res. Commun.	202 (3), 1308-1314	(1994)
MEDLINE	9433842		
PUBMED	7520237		
REFERENCE	2	(bases 1 to 957)	
AUTHORS	Stuyver, L., van Arnhem, W., Wyseur, A., Hernandez, F., Delaporte, E. and Maertens, G.		
TITLE	Classification of hepatitis C viruses based on phylogenetic analysis of the envelope 1 and nonstructural 5B regions and identification of five additional subtypes		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	91 (21), 10134-10138	(1994)
MEDLINE	9502399		
PUBMED	7524083		
COMMENT	Original source text: Hepatitis C virus (individual isolate BE95)		
FEATURES	CDNA to genomic RNA.		
source	Location/Qualifiers		
	1..957		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/isolate="BE95"		
	/db_xref="taxon:11103"		
	<1..>957		
	/codon_start=1		
	/product="precursor protein"		
	/protein_id="AAA65788.1"		
	/db_xref="GI:793911"		
	/translation="MSTNPKPQKTKRNTNRRPDVKKFGGQIVGVYLLPRGPRM		
	GVRAFKTSRQPRGRQPIPKARQPTSRGQPGYWPFLYANGLWAGWLLSPRG		
	SRNMGPNDRKSNRLKVIDTLTCGFADLMGYIPLVGGPIGGVARALAHGVRLD		
	GVNATGNLPGCSFIFILALLSLTVPASAVPYRNAGSIYHVTVNDPNSIIVYEDN		
	LILHAPGCVPCVMITNVSRQVITFTLSAPSLGAVTAPLRADVLAGGAALCSALY		
	VGDGCAFLVGMFTYRPRQATVQNCNCSIIYSHVTVGRMA"		
	1..573		
	/product="core protein"		
mat_peptide	574..957		
mat_peptide	/gene="E1"		
	/product="envelope protein"		
gene	574..957		
	/gene="E1"		
BASE COUNT	189 a 283 c 270 g 215 t		
ORIGIN			
	Query Match	95.8%; Score 602; DB 14; Length 957;	
	Best Local Similarity	100.0%; Pred. No. 2.3e-162; Indels 0; Gaps 0;	
	Matches	602; Conservative 0; Mismatches 0;	
QY	2	TGGGTAAGGTCATCGATACCCCTAAACGTTGGGGATTCGCCGATCTCATGGGGTATATCCCGC	61
Db	356	TGGGTAAGGTCATCGATACCCCTAAACGTTGGGGATTCGCCGATCTCATGGGGTATATCCCGC	415
QY	62	TCGTAGCGCGCCCATTTGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCTTG	121
Db	416	TCGTAGCGCGCCCATTTGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCTCTTG	475
QY	122	AGGACGGGGTAAACTATGCAACAGGGAATTTACCCGGTTGCTCTTTCTCTATCTTTATTC	181
Db	476	AGGACGGGGTAAACTATGCAACAGGGAATTTACCCGGTTGCTCTTTCTCTATCTTTATTC	535
QY	182	TTGCTCTTCTCTGCTGTCTGACCGTTCCGGCTCTGCGATTCCTACCGAAATGCCTCTG	241
Db	536	TTGCTCTTCTCTGCTGTCTGACCGTTCCGGCTCTGCGATTCCTACCGAAATGCCTCTG	595
QY	242	GGATTATCATGTTACCAATGATTTGCCAAACTCTTCCATAGTCTATGAGCAGATAACC	301
Db	596	GGATTATCATGTTACCAATGATTTGCCAAACTCTTCCATAGTCTATGAGCAGATAACC	655
QY	302	TGATCCTACAGCACCTGTTGCGTGTGCTGTGTGTCATGACAGGTAATGTGAGTAGTCT	361
Db	656	TGATCCTACAGCACCTGTTGCGTGTGCTGTGTGTCATGACAGGTAATGTGAGTAGTCT	715
QY	362	GGGTCCTAAATACCCCTACACTGTACGCCCGAGCGCTCGGAGCAGTCAACGCTCTCTTC	421
Db	716	GGGTCCTAAATACCCCTACACTGTACGCCCGAGCGCTCGGAGCAGTCAACGCTCTCTTC	775
QY	422	GGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCGTTATACGTAGGAG	481
Db	776	GGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCCCTCTGCTCCGCGTTATACGTAGGAG	835
QY	482	ACGCGTGTGGGCACTATTCTTGTAGGCCAAATGTTCACTATAGGCTCGCCAGCAG	541
Db	836	ACGCGTGTGGGCACTATTCTTGTAGGCCAAATGTTCACTATAGGCTCGCCAGCAG	895
QY	542	CTACGGTGAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG	601
Db	896	CTACGGTGAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG	955
QY	602	CA 603	
Db	956	CA 957	
RESULT 6			
LOCUS	A40649	Sequence 49 from Patent WO9425601.	
DEFINITION	A40649		
ACCESSION	A40649.1	GI:2296684	
VERSION			
KEYWORDS	unidentified		
SOURCE	unclassified		
ORGANISM	unclassified		
	1 (bases 1 to 959)		
REFERENCE	Maertens, G. and Stuyver, L.		
AUTHORS	NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES AND THEIR USE AS		
TITLE	THERAPEUTIC AND DIAGNOSTIC AGENTS		
JOURNAL	Patent: WO 9425601-A 49 10-NOV-1994;		
	INNOGENETICS NV (BE)		
COMMENT	Other publication CA 2139100 941110		
	Other publication AU 6722294 941121		
	Other publication CN 1108030 950906		
	Other publication FI 946066 941223		
	Other publication NO 944967 941221		
	Other publication JP 7508423T 950921.		

Db	898	CTACGGTGCAGAACTGCACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG	957		718	GGGTCCAAATTAACCCCTACACTGTCAGCCCGGAGCCTCGAGCAGTCAAGGCTCCTCTTC	777
Qy	602	CA 603			422	GGAGAGCCGTTGACTACCTAGCGGAGGCGCTCCCTCTGCTCCGCTTATAGTAGGAG	481
Db	958	CA 959			778	GGAGAGCCGTTGACTACCTAGCGGAGGCGCTCCCTCTGCTCCGCTTATAGTAGGAG	837
RESULT 8					482	ACGGTGTGGGCACTATTCTTGGTAGGCCAAATGTTACCTATAGGCTTCGCCAGCAG	541
AX031897					838	ACGGTGTGGGCACTATTCTTGGTAGGCCAAATGTTACCTATAGGCTTCGCCAGCAG	897
LOCUS	AX031897				542	CTACGGTGCAGAACTGCACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG	601
DEFINITION	Sequence 49 from Patent EP0984068.				898	CTACGGTGCAGAACTGCACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATGG	957
ACCESSION	AX031897				602	CA 603	
VERSION	AX031897.1	GI:10279047			958	CA 959	
KEYWORDS	unidentified						
SOURCE	unidentified						
ORGANISM	unclassified.						
REFERENCE	1						
AUTHORS	Maertens, G. and Stuyver, L.						
TITLE	Sequences of hepatitis c virus genotypes and their use as						
JOURNAL	therapeutic and diagnostic agents						
INNOGENETICS	Patent: EP 0984068-A 49 08-MAR-2000;						
LOCATION/Qualifiers	INNOGENETICS NV (BE)						
1.	.959						
/organism="unidentified"							
/mol type="genomic DNA"							
/db xref="taxon:32644"							
/clone="PC-3-4"							
3.	.>959						
/note="unnamed protein product"							
/codon start=1							
/protein id="CAC09766.1"							
/db xref="GI:10279048"							
/translation="MSTNPKPQKTRNTNRRPQDVKFPGGQIVGGVLLPRRPRM							
SVRAIKTSERSQRRRQPIPKARQPTGRSMQPGYPWPLYANGLWAGWLLSPRG							
SRPNWGNDRPKRSNLGKVIDTLTCGFADLMGYIPLVGGPIGGVARALAHGVRLVD							
GVNATGNLPCFSIFILALLSCLTVPASAVPRNAGSIYHVTNDCPNSSIVYEADN							
LILHAPGCPVCMTGNVSRCWQITPTLSAPSLGAVTAPLRRAVDYLAGAALCSALY							
VGDACGALFLVQMFTYRPHQATVQNCNCSIYSHVTHGRMA"							
BASE COUNT	189 a	285 c	270 g	215 t			
ORIGIN							
Query Match	95.6%;	Score 602;	DB 6;	Length 959;			
Best Local Similarity	100.0%;	Pred. No. 2.3e-162;					
Matches 602;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
0;							
Qy	2	TGGGTAAGGTGTCATCGATACCCCTAACCTGCGGATTCGCCGATCTCATGGGGTATATCCCGC	61		358	TGGGTAAGGTGTCATCGATACCCCTAACCTGCGGATTCGCCGATCTCATGGGGTATATCCCGC	417
Db	358	TGGGTAAGGTGTCATCGATACCCCTAACCTGCGGATTCGCCGATCTCATGGGGTATATCCCGC	417		62	TCGTAGGCGGCCCATTTGGGGCGCTCGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	121
Qy	62	TCGTAGGCGGCCCATTTGGGGCGCTCGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	121		418	TCGTAGGCGGCCCATTTGGGGCGCTCGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	477
Db	418	TCGTAGGCGGCCCATTTGGGGCGCTCGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	477		122	AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTCTTCTCTATCTTTATTC	181
Qy	122	AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTCTTCTCTATCTTTATTC	181		478	AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTCTTCTCTATCTTTATTC	537
Db	478	AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTCTTCTCTATCTTTATTC	537		182	TTGCTCTTCTCTGCTGTGACCGTTCGGGCTCTGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	241
Qy	182	TTGCTCTTCTCTGCTGTGACCGTTCGGGCTCTGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	241		538	TTGCTCTTCTCTGCTGTGACCGTTCGGGCTCTGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	597
Db	538	TTGCTCTTCTCTGCTGTGACCGTTCGGGCTCTGCAAGGGCTCTCGCACAGCGTGTAGGGTCTCTTG	597		242	GGATTTATCATGTTACATGATGCGCAACTCTCTCCATGCTCTATGAGGCGAGATAACC	301
Qy	242	GGATTTATCATGTTACATGATGCGCAACTCTCTCCATGCTCTATGAGGCGAGATAACC	301		598	GGATTTATCATGTTACATGATGCGCAACTCTCTCCATGCTCTATGAGGCGAGATAACC	657
Db	598	GGATTTATCATGTTACATGATGCGCAACTCTCTCCATGCTCTATGAGGCGAGATAACC	657		302	TGATCTCTACAGCACCTGGTGTGGTGTGCTGTGCTATGACAGGTAAATGTGAGTAGATGCT	361
Qy	302	TGATCTCTACAGCACCTGGTGTGGTGTGCTGTGCTATGACAGGTAAATGTGAGTAGATGCT	361		658	TGATCTCTACAGCACCTGGTGTGGTGTGCTGTGCTATGACAGGTAAATGTGAGTAGATGCT	717
Db	658	TGATCTCTACAGCACCTGGTGTGGTGTGCTGTGCTATGACAGGTAAATGTGAGTAGATGCT	717		362	GGGTCCAAATTAACCCCTACACTGTCAGCCCGGAGCCTCGGAGCAGTCAAGGCTCCTCTTC	421
Qy	362	GGGTCCAAATTAACCCCTACACTGTCAGCCCGGAGCCTCGGAGCAGTCAAGGCTCCTCTTC	421				

RESULT 12	AX031631	LOCUS	AX031631	53 bp	DNA	linear	PAT 20-SEP-2000
			sequence	from Patent	EP1004670.		

```
|||||
898 CTACGGTGCAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATCG 957
|||||
QY 602 CA 603
||
Db 958 CA 959

RESULT 13
AX031901
LOCUS AX031901 959 bp DNA linear PAT 20-SEP-2000
DEFINITION Sequence 53 from Patent EP0984068.
ACCESSION AX031901
VERSION AX031901.1 GI:10279051
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
1 Maertens, G. and Stuyver, L.
AUTHORS Sequences of hepatitis C virus genotypes and their use as
TITLE therapeutic and diagnostic agents
JOURNAL Patent: EP 0984068-A 53 08-MAR-2000;
INNOGENETICS NV (BE)
FEATURES
source Location/Qualifiers
1 .959
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/clone="PC C/E1"
3..959
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09768.1"
/db_xref="GI:10279052"
/translation="MSTNPKPQKTKNTNRRPQDKVFPGGQIVGGVLLPRRPRGM
GVATRKTSERSQPRRQPIPKARQPTGRSGWQPGYPMPLVANEGLWAGWLLSPRG
SRPNPNDPRKSRNLGKVIDLTLCGFADLMGYIPLVGGPXXGVARALAHGVRLD
GVNYATGNLPGCSFSIFILALLSLCTVPASAVPYRNASGIYHVTNDPCNSSIYVEADN
LILHAPGCVPCVMTGNVSRWCVOITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALY
VGDACGALPLVQMFTYRPRQATVQNCNCSISYGHVTHRMA"
BASE COUNT 188 a 285 c 269 g 214 t 3 others
ORIGIN
Query Match 95.4%; Score 600.8; DB 6; Length 959;
Best Local Similarity 99.5%; Pred. No. 5.1e-162;
Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 61
|||||
Db 358 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 417
|||||
QY 62 TCGTAGCGCGGCCCTTTGGGGCGTCGCAAGGGCTCTCGCACACGGGTGTAGGGTCTCTTG 121
|||||
Db 418 TCGTAGCGCGGCCCTTTGGGGCGTCGCAAGGGCTCTCGCACACGGGTGTAGGGTCTCTTG 477
|||||
QY 122 AGGACGGGGTAAACTATGCAACAGGAAATTTACCGGTTGCTCTTCTCTATCTTTATTC 181
|||||
Db 478 AGGACGGGGTAAACTATGCAACAGGAAATTTACCGGTTGCTCTTCTCTATCTTTATTC 537
|||||
QY 182 TTGCTCTTCTCGTCTGACCGTTCCGGCCTCTGACGTTCCCTACCGAAATGCCTCTG 241
|||||
Db 538 TTGCTCTTCTCGTCTGACCGTTCCGGCCTCTGACGTTCCCTACCGAAATGCCTCTG 597
|||||
QY 242 GGAATTATCATGTTACCAATGATGCGCAAACTCTTCATAGTCTATGAGGCGATGAACC 301
|||||
Db 598 GGAATTATCATGTTACCAATGATGCGCAAACTCTTCATAGTCTATGAGGCGATGAACC 657
|||||
QY 302 TGATCTCTACGACCTGTTGCGTCTGTGTCATGACAGTAATGTAGTAGTAGTCT 361
|||||
Db 658 TGATCTCTACGACCTGTTGCGTCTGTGTCATGACAGTAATGTAGTAGTAGTCT 717
|||||
QY 362 GGGTCCAAATTAACCTACACTGTCTAGCCCCCGAGCCTCGAGGAGCAGTCAAGGCTCTCTTC 421
|||||

Db 718 GGGTCCAAATTAACCTACACTGTCTAGCCCCCGAGCCTCGAGGAGCAGTCAAGGCTCTCTTC 777
|||||
QY 422 GGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGTCTCGGTTTATAGTAGGAG 481
|||||
Db 778 GGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTCTGTCTCGGTTTATAGTAGGAG 837
|||||
QY 482 ACGGTGTGGGCGACATATTTCTTGGTAGGCCAAATGTTACCTATAGGCTCTGCCACACG 541
|||||
Db 838 ACGGTGTGGGCGACATATTTCTTGGTAGGCCAAATGTTACCTATAGGCTCTGCCACACG 897
|||||
QY 542 CTACGGTGCAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATCG 601
|||||
Db 898 CTACGGTGCAGAACTGCAACTGTTCCATTACAGTGGCCATGTTACCGGCCACCGGATCG 957
|||||

QY 602 CA 603
||
Db 958 CA 959

RESULT 14
AX032171
LOCUS AX032171 959 bp DNA linear PAT 20-SEP-2000
DEFINITION Sequence 53 from Patent EP0984067.
ACCESSION AX032171
VERSION AX032171.1 GI:10279234
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
1 Maertens, G. and Stuyver, L.
AUTHORS Sequences of hepatitis C virus genotypes and their use as
TITLE therapeutic and diagnostic agents
JOURNAL Patent: EP 0984067-A 53 08-MAR-2000;
INNOGENETICS NV (BE)
FEATURES
source Location/Qualifiers
1 .959
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/clone="PC C/E1"
3..959
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09844.1"
/db_xref="GI:10279235"
/translation="MSTNPKPQKTKNTNRRPQDKVFPGGQIVGGVLLPRRPRGM
GVATRKTSERSQPRRQPIPKARQPTGRSGWQPGYPMPLVANEGLWAGWLLSPRG
SRPNPNDPRKSRNLGKVIDLTLCGFADLMGYIPLVGGPXXGVARALAHGVRLD
GVNYATGNLPGCSFSIFILALLSLCTVPASAVPYRNASGIYHVTNDPCNSSIYVEADN
LILHAPGCVPCVMTGNVSRWCVOITPTLSAPSLGAVTAPLRRAVDYLAGGAALCSALY
VGDACGALPLVQMFTYRPRQATVQNCNCSISYGHVTHRMA"
BASE COUNT 188 a 285 c 269 g 214 t 3 others
ORIGIN
Query Match 95.4%; Score 600.8; DB 6; Length 959;
Best Local Similarity 99.5%; Pred. No. 5.1e-162;
Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 61
|||||
Db 358 TGGGTAAGGTCTATCGATACCTAAACGTGCGGATTCGCCGATCTCATGGGGTATATCCCGC 417
|||||
QY 62 TCGTAGCGCGGCCCTTTGGGGCGTCGCAAGGGCTCTCGCACACGGGTGTAGGGTCTCTTG 121
|||||
Db 418 TCGTAGCGCGGCCCTTTGGGGCGTCGCAAGGGCTCTCGCACACGGGTGTAGGGTCTCTTG 477
|||||
QY 122 AGGACGGGGTAAACTATGCAACAGGAAATTTACCGGTTGCTCTTCTCTATCTTTATTC 181
|||||
Db 478 AGGACGGGGTAAACTATGCAACAGGAAATTTACCGGTTGCTCTTCTCTATCTTTATTC 537
|||||
QY 182 TTGCTCTTCTCGTCTGACCGTTCCGGCCTCTGACGTTCCCTACCGAAATGCCTCTG 241
|||||
QY 242 GGAATTATCATGTTACCAATGATGCGCAAACTCTTCATAGTCTATGAGGCGATGAACC 301
|||||
Db 598 GGAATTATCATGTTACCAATGATGCGCAAACTCTTCATAGTCTATGAGGCGATGAACC 657
|||||
QY 302 TGATCTCTACGACCTGTTGCGTCTGTGTCATGACAGTAATGTAGTAGTAGTCT 361
|||||
Db 658 TGATCTCTACGACCTGTTGCGTCTGTGTCATGACAGTAATGTAGTAGTAGTCT 717
|||||
QY 362 GGGTCCAAATTAACCTACACTGTCTAGCCCCCGAGCCTCGAGGAGCAGTCAAGGCTCTCTTC 421
|||||
```

Db	538																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 16:53:58 ; Search time 174.169 Seconds
(without alignments)
9764.351 Million cell updates/sec

Title: US-09-899-303A-31

Perfect score: 630

Sequence: 1 ATGGTAGGTCATCGATAC.....TCATGATGACTGGTAATAG 630

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	100.0	630	17	AAT12966 HCV El construct H
2	630	100.0	630	24	AAQ48930 Hepatitis C virus
3	602	95.6	959	15	AAQ78047 Hepatitis C virus
4	600.8	95.4	959	15	AAQ78049 Hepatitis C virus
5	597.2	94.8	959	15	AAQ78048 Hepatitis C virus
6	580	92.1	580	15	AAQ78045 Hepatitis C virus
7	580	92.1	580	15	AAQ78046 Hepatitis C virus
8	573.2	91.0	579	15	AAQ78092 Hepatitis C virus

9	510.2	81.0	579	15	AAQ78113 Hepatitis C virus
10	510.2	81.0	579	15	AAQ78093 Hepatitis C virus
11	347.2	55.1	576	16	AAQ83891 Hepatitis C virus
12	347.2	55.1	576	17	AAT16605 Hepatitis C virus
13	344	54.6	576	16	AAQ83893 Hepatitis C virus
14	344	54.6	576	17	AAT16607 Hepatitis C virus
15	340.8	54.1	576	16	AAQ83892 Hepatitis C virus
16	340.8	54.1	576	17	AAT16606 Hepatitis C virus
17	337.6	53.6	576	16	AAQ83894 Hepatitis C virus
18	337.6	53.6	576	16	AAQ83890 Hepatitis C virus
19	337.6	53.6	576	17	AAT16604 Hepatitis C virus
20	337.6	53.6	576	17	AAT16608 Hepatitis C virus
21	335.6	53.3	630	17	AAT12965 HCV El construct H
22	335.6	53.3	630	24	AAQ48929 Hepatitis C virus
23	331.6	52.6	9595	20	AAQ24843 Infectious hepatitis
24	331.6	52.6	9595	22	AAQ86939 Nucleotide sequenc
25	331.6	52.6	9595	22	AAF23492 Infectious Hepatit
26	331.6	52.6	9599	20	AAQ24833 Infectious hepatitis
27	331.2	52.6	576	16	AAQ83889 Hepatitis C virus
28	331.2	52.6	576	17	AAT16603 Hepatitis C virus
29	326.8	51.9	2187	19	ABA03491 Cuticle protein 1
30	326.8	51.9	2540	14	AAQ43889 NANBHV genomic fra
31	326.8	51.9	2540	15	AAQ63753 Fragment of NANB h
32	325.2	51.6	1863	12	AAQ15363 NANB hepatitis vir
33	325.2	51.6	1880	13	AAQ24467 NANB hepatitis C vir
34	325.2	51.6	2540	13	AAQ29628 HCV El construct H
35	324.8	51.6	795	17	AAT12705 Hepatitis C virus
36	324.8	51.6	795	24	AAQ48914 Hepatitis C virus
37	324.8	51.6	2082	24	AAQ48939 Hepatitis C virus
38	324.8	51.6	2086	17	AAT12973 HCV El construct H
39	324.8	51.6	2229	19	ABA03487 Cuticle protein 1
40	324.8	51.6	2433	17	AAT12974 HCV El construct H
41	323.6	51.4	2116	12	AAQ12242 Encodes PT-NANBH v
42	323.6	51.4	2187	19	ABA03492 Cuticle protein 1
43	323.4	51.3	633	17	AAT12706 HCV El construct H
44	323.4	51.3	633	24	AAQ48915 Hepatitis C virus
45	323.2	51.3	9472	14	AAQ33282 Korean hepatitis C

ALIGNMENTS

RESULT 1

AAT12966

ID AAT12966 standard; DNA; 630 BP.

XX AAT12966;

XX AAT12966;

XX 24-SEP-1996 (first entry)

XX HCV El construct HCC163.

DE HCV El construct HCC163.

XX HCV; E1, E2, disulphide bond cleavage; envelope protein; vaccine; human;

XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX ss.

XX Hepatitis C virus.

OS Hepatitis C virus.

XX WO9604385-A2.

PN 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

PA Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI; 1996-129401/13.

DR Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

XX

PT

AAQ78049	DB	418	TCGTAGGCGGCCCRRTTGGGGCGTCGCAAGGGCTCTCCACACGGGTGTGAGGGTCTTTC	477
ID AAQ78049 standard; cDNA; 959 BP.				
XX AC				
XX AAQ78049;				
XX 25-MAR-2003 (updated)				
DT 02-AUG-1995 (first entry)				
XX				
DE Hepatitis C virus E1 region.				
XX				
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;				
KW classification; immunisation; prophylaxis; serotyping; ss.				
KW				
XX				
OS Hepatitis C virus type 5a.				
XX				
XX Key Location/Qualifiers				
FT CDS 3..959				
FT /*tag= a				
FT /product= E1 polypeptide.				
XX				
XX WO9425601-A2.				
XX				
XX 10-NOV-1994.				
XX				
XX 27-APR-1994; 94WO-EP01323.				
XX				
XX 27-APR-1993; 93EP-0401099.				
PR 05-AUG-1993; 93EP-0402019.				
XX				
XX (INNO-) INNOGENETICS NV SA.				
XX				
XX Maertens G, Stuyver L;				
XX				
XX WPI: 1994-358277/44.				
XX P-PSDB; AAR63297.				
XX				
XX New polynucleotide sequences from hepatitis C virus - and related				
PT vectors, polypeptide(s) and antibodies, useful for immunisation,				
PT treatment, diagnosis and typing of HCV isolates				
XX				
XX Claim 3; Page 144; 404pp; English.				
XX				
XX Compositions comprising at least 5, and pref. 8 or more contiguous				
CC nucleotides selected from an HCV type 3 genomic sequence, more				
CC particularly (i) the region spanning positions 417-957 of the				
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions				
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning				
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the				
CC region spanning positions 8023-8235 of the NS5 region of the BR36				
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic				
CC sequence, or, from a subtype 2d genomic sequence, a type 4 genomic				
CC sequence; or the coding region of subtype 5a, may be used as primers				
CC to amplify nucleic acid from an isolate belonging to a specific				
CC genotype, or as a probe for specific detection/classification of				
CC nucleic acid. Polypeptides encoded by the nucleotides in such				
CC compositions may be used for immunisation against HCV, for the				
CC detection of antibodies directed against HCV and for serotyping.				
CC This sequence corresponds to the E1 region of HCV subtype 5a and				
CC is taken from a clone designated PC C/E1.				
CC (Updated on 25-MAR-2003 to correct PN field.)				
XX				
XX Sequence 959 BP; 188 A; 285 C; 269 G; 214 T; 3 other;				
SQ				
Query Match 95.4%; Score 600.8; DB 15; Length 959;				
Best Local Similarity 99.5%; Pred. No. 4.6e-177;				
Matches 599; Conservative 3; Mismatches 0; Indels 0; Gaps 0;				
QY 2 TGGGTAAAGTCAATCGATACCCCTAACGTCGGGATTCGCGATCTCATGGGTATATCCGCG 61				
DB 358 TGGGTAAAGTCAATCGATACCCCTAACGTCGGGATTCGCGATCTCATGGGTATATCCGCG 417				
QY 62 TCGTAGCGGCCCATTTGGGGGCGTCGAAGGGCTCTCCACACGGTGTGAGGGTCTTTC 121				

DR WPI: 1994-358277/44.
DR P-PSDB; AAR63296.

XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates

XX Claim 3; Page 141-142; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
CC sequence, or, from a subtype 2d genomic sequence, a type 4 genomic
CC sequence; or the coding region of subtype 5a, may be used as primers
CC to amplify nucleic acid from an isolate belonging to a specific
CC genotype, or as a probe for specific detection/classification of
CC nucleic acid. Polypeptides encoded by the nucleotides in such
CC compositions may be used for immunisation against HCV, for the
CC detection of antibodies directed against HCV and for serotyping.
CC This sequence corresponds to the E1 region of HCV subtype 5a and
CC is taken from a clone designated PC-3-8.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 959 BP; 188 A; 287 C; 270 G; 214 T; 0 other;

Query Match 94.8%; Score 597.2; DB 15; Length 959;
Best Local Similarity 99.5%; Pred. No. 6.1e-176;
Matches 599; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TGGGTAAGGTATCGATACCTTAACGTGCGGATTGCGCGATCTCATGGGGTATATCCCGC 61
DB 358 TGGGTAAGGTATCGATACCTTAACGTGCGGATTGCGCGATCTCATGGGGTATATCCCGC 417

QY 62 TCGTAGGGCGGCCCATTTGGGGGCGTGCAGAGGCTCTCGCACACGGTGTGAGGGTCTTGG 121
DB 418 TCGTAGGGCGGCCCATTTGGGGGCGTGCAGAGGCTCTCGCACACGGTGTGAGGGTCTTGG 477

QY 122 AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTTTCTCTATCTTTATTC 181
DB 478 AGGACGGGTAACTATGCAACAGGGAATTTACCGGTTGCTTTCTCTATCTTTATTC 537

QY 182 TTGCTCTTCTCTGCTGCTGACCGTTTCGGGCTCTGCAAGTTCCTACCGAAATGCTCTG 241
DB 538 TTGCTCTTCTCTGCTGCTGACCGTTTCGGGCTCTGCAAGTTCCTACCGAAATGCTCTG 597

QY 242 GGATTTATCATGTTACCAATGATTGGCCAAACTCTTCCATAGTCTATGAGGAGATAACC 301
DB 598 GGATTTATCATGTTACCAATGATTGGCCAAACTCTTCCATAGTCTATGAGGAGATAACC 657

QY 302 TGATCTTACACGACCTGTTGGTGGCTTGTGTGTCATGACAGTAATGTAGTAGTGTCT 361
DB 658 TGATCTTACACGACCTGTTGGTGGCTTGTGTGTCATGACAGTAATGTAGTAGTGTCT 717

QY 362 GGGTCCAAATTAACCTTACACTGTACGCTCGAGCTCGAGCAGTACGGTCTCTCTTC 421
DB 718 GGGTCCAAATTAACCTTACACTGTACGCTCGAGCTCGAGCAGTACGGTCTCTCTTC 777

QY 422 GGAGAGCCGTTGACTACCTAGCGGGAGGGCTGCGCTCTGCTCCGGTTATAGTAGGAG 481
DB 778 GGAGAGCCGTTGACTACCTAGCGGGAGGGCTGCGCTCTGCTCCGGTTATAGTAGGAG 837

QY 482 ACGCGTGTGGGCACTATTCTTGGTAGGCCAAATGTTACCTATATAGGCTCTGCCAGCAG 541
DB 838 ACGCGTGTGGGCACTATTCTTGGTAGGCCAAATGTTACCTATATAGGCTCTGCCAGCAG 897

QY 542 CTACGGTGCAGAACTGCACTGTTTCCATTACAGTGCCCATGTTACCGGCCACCGATGG 601
DB 898 CTACGGTGCAGAACTGCACTGTTTCCATTACAGTGCCCATGTTACCGGCCACCGATGG 957

QY 602 CA 603
DB 958 CA 959

RESULT 6
AAQ78045
ID AAQ78045 standard; cDNA; 580 BP.
XX AAQ78045;
XX 25-MAR-2003 (updated)
DT 02-AUG-1995 (first entry)
XX Hepatitis C virus E1 region.
XX Hepatitis C virus type 5a.
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping; ss.

XX Hepatitis C virus type 5a.
FH Key Location/Qualifiers
CDS 2..580
FT /*tag= a
FT /product= E1 polypeptide.
XX W09425601-A2.
XX 10-NOV-1994.
XX 27-APR-1994; 94WO-EP01323.
XX 27-APR-1993; 93EP-0401099.
XX 05-AUG-1993; 93EP-0402019.
XX (INNO-) INNOGENETICS NV SA.
XX Maertens G, Stuyver L;
WPI: 1994-358277/44.
P-PSDB; AAR63293.

New polynucleotide sequences from hepatitis C virus - and related
vectors, polypeptide(s) and antibodies, useful for immunisation,
treatment, diagnosis and typing of HCV isolates

Claim 3; Page 133-134; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the
CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions
CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning
CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the
CC region spanning positions 8023-8235 of the NS5 region of the BR36
CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic
CC sequence, or, from a subtype 2d genomic sequence, a type 4 genomic
CC sequence; or the coding region of subtype 5a, may be used as primers
CC to amplify nucleic acid from an isolate belonging to a specific
CC genotype, or as a probe for specific detection/classification of
CC nucleic acid. Polypeptides encoded by the nucleotides in such
CC compositions may be used for immunisation against HCV, for the
CC detection of antibodies directed against HCV and for serotyping.
CC This sequence corresponds to the E1 region of HCV subtype 5a and
CC is taken from a clone designated PC-4-1.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 580 BP; 111 A; 165 C; 155 G; 149 T; 0 other;

Query Match 92.1%; Score 580; DB 15; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.1e-170;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 564 TTCCATTTACAGTGGCCATGTTACCGGCCACCGGATGGCA 603
 Db 541 TTCCATTTACAGTGGCCATGTTACCGGCCACCGGATGGCA 580

RESULT 8

AAQ78092
 ID AAQ78092 standard; cDNA; 579 BP.

AC AAQ78092;

DT 25-MAR-2003 (updated)

DT 15-AUG-1995 (first entry)

DE Hepatitis C virus E1 region.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping; ss.

OS Hepatitis C virus.

FH Key Location/Qualifiers

FT CDS 1..579

FT /*tag= a

FT /product= E1 polypeptide.

FT mat_peptide 1..576

FT /*tag= b

XX WO9425601-A2.

PN 10-NOV-1994.

XX 27-APR-1994; 94WO-BP01323.

XX 27-APR-1993; 93EP-0401099.

PR 05-AUG-1993; 93EP-0402019.

XX (INNO-) INNOGENETICS NV SA.

PA Maertens G, Stuyver L;

XX MPI: 1994-358277/44.

DR P-PSDB; AAR63354.

XX New polynucleotide sequences from hepatitis C virus - and related

PT vectors, polypeptide(s) and antibodies, useful for immunisation,

PT treatment, diagnosis and typing of HCV isolates

XX Claim 3; Page 202-203; 404pp; English.

PS Compositions comprising at least 5, and pref. 8 or more contiguous

XX nucleotides selected from an HCV type 3 genomic sequence, more

CC particularly (i) the region spanning positions 417-957 of the

CC Core/E1 region of HCV subtype 3a; (ii) the region spanning positions

CC 4664-4730 of the NS3 region of HCV type 3; (iii) the region spanning

CC positions 4892-5292 of the NS3/4 region of HCV type 3; (iv) the

CC region spanning positions 8023-8235 of the NS5 region of the BR36

CC subgroup of HCV subtype 3a; or (v) an HCV subtype 3c genomic

CC sequence, may be used as primers to amplify nucleic acid from an

CC isolate belonging to a specific genotype, or as a probe for specific

CC detection/classification of nucleic acid. Polypeptides encoded by

CC the nucleotides in such compositions may be used for immunisation

CC against HCV, for the detection of antibodies directed against HCV

CC and for serotyping. This sequence corresponds to the E1 region

CC of HCV.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 579 BP; 108 A; 167 C; 156 G; 148 T; 0 other;

XX Query Match 91.0%; Score 573.2; DB 15; Length 579;

XX Best Local Similarity 99.5%; Pred. No. 1.5e-168;

XX Matches 575; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 25 ACGTGGGATTCGCGATCTCATGGGGTATATCCCGCTCGTAGGGCGGCCCATTTGGGGGC 84
 Db 1 ACGTGGGATTCGCGATCTCATGGGGTACATCCCGCTCGTAGGGCGGCCCATTTGGGGGC 60
 QY 85 GTCCGAAGGGCTCTCGCACACGGTGTGAGGTCCTTGAGGAGGGGTAACTATGCAACA 144
 Db 61 GTCCGAAGGGCTCTCGCACACGGTGTGAGGTCCTTGAGGAGGGGTAACTATGCAACA 120
 QY 145 GGGAAATTTACCGGTTGCTCTTTCTATCTTTATTTCTTGTCTCTTCTCGTGTCTGACC 204
 Db 121 GGGAAATTTACCGGTTGCTCTTTCTATCTTTATTTCTTGTCTCTTCTCGTGTCTGACC 180
 QY 205 GTTCGGGCTCTGCGAGTTCCTTACCGAAATGCTCTGGGATTTATCATGTTACCAATGAT 264
 Db 181 GTTCGGGCTCTGCGAGTTCCTTACCGAAATGCTCTGGGATTTATCATGTTACCAATGAT 240
 QY 265 TGCCAAACTCTTTCATAGTCTATGAGGAGAGATAACCTGATCTACACGACCTGGTTGC 324
 Db 241 TGCCAAACTCTTTCATAGTCTATGAGGAGAGATAACCTGATCTACACGACCTGGTTGC 300
 QY 325 GTGCTTGTTCATGACAGGTAATGTGAGTAGATGCTGGGTCCAAATTAACCCCTACACTG 384
 Db 301 GTGCTTGTTCATGACAGGTAATGTGAGTAGATGCTGGGTCCAAATTAACCCCTACACTG 360
 QY 385 TCAGCCCCGAGCTCGGAGCAGTCACGGCTCTCTTCGGAGAGCCGTTGACTACTAGCG 444
 Db 361 TCAGCCCCGAGCTCGGAGCAGTCACGGCTCTCTTCGGAGAGCCGTTGACTACTAGCG 420
 QY 445 GGAGGGGCTGCCCTCTGTCTCCGTTATACGTAGGAGAGCGGTGGGGCAGCTATTCTTGG 504
 Db 421 GGAGGGGCTGCCCTCTGTCTCCGTTATACGTAGGAGAGCGGTGGGGCAGCTATTCTTGG 480
 QY 505 GTAGGCCAAATGTTTACCTATAGGCTCCGCCAGCAGCTACGGTGCAGAACTGCAACTGT 564
 Db 481 GTAGGCCAAATGTTTACCTATAGGCTCCGCCAGCAGCTACGGTGCAGAACTGCAACTGT 540
 QY 565 TCATTTACAGTGGCCATGTTACCGGCCACCGGATGGC 602
 Db 541 TCATTTACAGTGGCCATGTTACCGGCCACCGGATGGC 578

RESULT 9

AAQ78113
 ID AAQ78113 standard; cDNA; 579 BP.

AC AAQ78113;

XX 25-MAR-2003 (updated)

DT 18-AUG-1995 (first entry)

DE Hepatitis C virus NS5B region.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;

XX classification; immunisation; prophylaxis; serotyping; ss.

OS Hepatitis C virus.

FH Key Location/Qualifiers

FT CDS 1..579

FT /*tag= a

FT /product= NS5B polypeptide.

FT mat_peptide 1..576

FT /*tag= b

XX WO9425601-A2.

PN 10-NOV-1994.

XX 27-APR-1994; 94WO-BP01323.

XX 27-APR-1993; 93EP-0401099.

PR 05-AUG-1993; 93EP-0402019.

XX

PD 22-FEB-1996.
 XX 15-AUG-1995; 95WO-US10398.
 XX 15-AUG-1994; 94US-0290665.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (USSH) US SEC DEPT HEALTH.
 XX Bukh J, Miller RH, Purcell RH;
 XX WPI; 1996-139709/14.
 XX P-PSDB; AAR89551.
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins -
 PT used to determine HCV genotype and as vaccines against HCV infection
 XX Claim 1; Page 107; 340pp; English.
 XX AAT16559-T16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
 CC isolates. The isolated sequences are useful for the prodn. of primers
 CC useful for detecting the presence of HCV in a sample, the primers
 CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
 CC can be used in vaccines for immunising against HCV infection. The
 CC proteins may also be used to detect antibodies against HCV in serum,
 CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
 CC used in the prevention of HCV infection.
 XX Sequence 576 BP; 106 A; 161 C; 168 G; 141 T; 0 other;
 SQ
 Query Match 55.1%; Score 347.2; DB 17; Length 576;
 Best Local Similarity 90.7%; Pred. No. 5.3e-98;
 Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 220 GTTCCCTACCGAAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 279
 DB 1 GTCCCTACCGAAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 60
 QY 280 ATAGTCTATGAGGAGATTAACCTGATCTACACGACCTGGTTCGGTTCCTTGTGTCATG 339
 DB 61 ATAGTCTACGAGGCTGATAACCTGATCTGACGACCTGGTTCGGTTCCTTGTGTCAG 120
 QY 340 ACAGTAAATGAGTAGATGCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 399
 DB 121 GAAGGAGTACAGGCTCTCTTCGGAGGCTGCTGACTACTTACGCGGAGGGCTGCGCTC 180
 QY 400 GGAGCAGTACAGGCTCTCTTCGGAGGCTGCTGACTACTTACGCGGAGGGCTGCGCTC 459
 DB 181 GGAGGCTTACAGGCTCTCTTCGGAGGCTGCTGACTACTTACGCGGAGGGCTGCGCTC 240
 QY 460 TGCTCCGGCTTATACGTAGGAGACCGCTGTCGGGACCTTCTTGGTAGGCCAATGTC 519
 DB 241 TGCTCCGCACTATAGCTGGGACCGCTGTCGGGACCTTCTTGGTAGGCCAATGTC 300
 QY 520 ACCTATAGGCTCGCCAGCAGCTACGCTGCGGAGGCTGCTGCTTCCATTTACAGTGGC 579
 DB 301 ACCTATAGGCTCGCCAGCAGCTACGCTGCGGAGGCTGCTGCTTCCATTTACAGGCGC 360
 QY 580 CATGTTACCGGACCGGATGGCATGGGATGATGATGATGATGATGATGATGATGATGATG 627
 DB 361 CATATCACCGGACCGGATGGCATGGGATGATGATGATGATGATGATGATGATGATGATG 408
 RESULT 13
 ID AAQ83893
 XX AAQ83893 standard; cDNA; 576 BP.
 AC AAQ83893;
 XX 25-MAR-2003 (updated)
 DT 19-SEP-1995 (first entry)
 XX Hepatitis C virus envelope 1 gene cDNA isolate SA7.

XX Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate SA7;
 KW diagnosis; vaccines; antibodies; antisera; gene inhibition; sa.
 XX Hepatitis C virus.
 XX Key Location/Qualifiers
 FT mat_peptide 1..576
 FT /*tag= a
 XX WO9501442-A2.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US07320.
 XX 29-JUN-1993; 93US-0086428.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Bukh J, Miller RH, Purcell RH;
 XX WPI; 1995-061006/08.
 XX P-PSDB; AAR69682.
 XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
 PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
 PT and in vaccines
 XX Claim 1; Page 78; 186pp; English.
 XX AAQ83893 encodes AAR69682 hepatitis C virus (HCV) envelope 1 (E1)
 CC protein isolate SA7, both can be used for the diagnosis of HCV
 CC infection, and in the prodn. of anti-HCV vaccines, antibodies
 CC and antisera. The cDNA may also be used to inhibit the expression
 CC of the HCV E1 gene.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 576 BP; 104 A; 165 C; 168 G; 139 T; 0 other;
 SQ
 Query Match 54.6%; Score 344; DB 16; Length 576;
 Best Local Similarity 90.2%; Pred. No. 5.3e-97;
 Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 220 GTTCCCTACCGAAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 279
 DB 1 GTCCCTACCGAAATGCTCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 60
 QY 280 ATAGTCTATGAGGAGATTAACCTGATCTACACGACCTGGTTCGGTTCCTTGTGTCATG 339
 DB 61 ATAGTCTATGAGGCTGACAACTGATCTGCGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 120
 QY 340 ACAGGTAATGAGTAGATGCTGGGATTTATCATGTTACCAATGATGCCCAAACTCTTCC 399
 DB 121 CAAATTAATGTCAGTAGGCTGCTGGTTCACCCCAATGATGCCCAAACTCTTCC 180
 QY 400 GGAGCAGTACAGGCTCTCTTCGGAGGCTGCTGACTACTTACGCGGAGGGCTGCGCTC 459
 DB 181 GGAGGCTTACAGGCTCTCTTCGGAGGCTGCTGACTACTTACGCGGAGGGCTGCGCTC 240
 QY 460 TGCTCCGGCTTATACGTAGGAGACCGCTGTCGGGACCTTCTTGGTAGGCCAATGTC 519
 DB 241 TGCTCCGCACTATAGCTGGGACCGCTGTCGGGACCTTCTTGGTAGGCCAATGTC 300
 QY 520 ACCTATAGGCTCGCCAGCAGCTACGCTGCGGAGGCTGCTGCTTCCATTTACAGTGGC 579
 DB 301 AGCTATAGGCTCGCCAGCAGCTACGCTGCGGAGGCTGCTGCTTCCATTTACAGTGGC 360
 QY 580 CATGTTACCGGACCGGATGGCATGGGATGATGATGATGATGATGATGATGATGATGATG 627
 DB 361 CATATCACCGGACCGGATGGCATGGGATGATGATGATGATGATGATGATGATGATGATG 408

```
RESULT 14
AAT16607
ID AAT16607 standard; cDNA; 576 BP.
XX
AC AAT16607;
XX
DT 30-SEP-1996 (first entry)
XX
DE Hepatitis C virus isolate SA7 envelope 1 gene.
XX
KW HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;
KW hepatitis; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1..576
FT /tag= a
FT /product= envelope-1_protein
FT /note= "does not contain start or stop codon"
XX
PN W09605315-A2.
XX
PD 22-FEB-1996.
XX
PF 15-AUG-1995; 95WO-US10398.
XX
PR 15-AUG-1994; 94US-0290665.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
DR WPI; 1996-139709/14.
XX
DR P-PSDB; AAR89553.
XX
XX
PT DNA and amino acid sequence of HCV envelope 1 and core proteins -
PT used to determine HCV genotype and as vaccines against HCV infection
XX
PS Claim 1; Page 108; 340pp; English.
XX
CC AAT16559-T16609 are cDNAs encoding the E1 (envelope-1) protein of 51 HCV
CC isolates. The isolated sequences are useful for the prodn. of primers
CC useful for detecting the presence of HCV in a sample, the primers
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs
CC can be used in vaccines for immunising against HCV infection. The
CC proteins may also be used to detect antibodies against HCV in serum,
CC saliva, lymphocytes or other mononuclear cells. The antibodies may be
CC used in the prevention of HCV infection.
XX
SQ Sequence 576 BP; 104 A; 165 C; 168 G; 139 T; 0 other;

Query Match 54.6%; Score 344; DB 17; Length 576;
Best Local Similarity 90.2%; Pred. No. 5.3e-97;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATGCCTCTGGGATTATCATGTTACCAATGATGCGCCAAACTCTTCC 279
DB 1 GTCCCTACCGAATGCCTCTGGGTTTATCATGTACCAATGATGCGCCAAACTCTTCC 60
QY 280 ATAGTCTATGAGCAGATAACCTGATCTACAGCACCTGGTTGCGTTCGTGTGTCATG 339
DB 61 ATAGTCTATGAGGCTGACAACTGATCTCTGACGACCTGGTTGCGTGTGTCATG 120
QY 340 ACAGGTAATGTAGTAGTAGTCTGGTCCAAATTAACCCCTACACTGTACAGCCCGAGCCTC 399
DB 121 CAAAATAATGTACGTAGTAGTCTGGGTCCAAATCACCCCAACATTTGTACAGCCCGAACCTC 180
QY 400 GGAGCAGTACGGCTCCTCTTCGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCCCTC 459
DB 181 GGAGCGGTACGGCTCCTCTTCGGAGAGCGGTTGACTACCTAGCGGAGGGGCTGCCCTC 240

RESULT 15
AAQ83892
ID AAQ83892 standard; cDNA; 576 BP.
XX
AC AAQ83892;
XX
DT 25-MAR-2003 (updated)
DT 19-SEP-1995 (first entry)
XX
DE Hepatitis C virus envelope 1 gene cDNA isolate SA6.
XX
KW Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate SA6;
KW diagnosis; vaccines; antibodies; antisera; gene inhibition; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..576
FT /tag= a
XX
PN W09501442-A2.
XX
PD 12-JAN-1995.
XX
PF 28-JUN-1994; 94WO-US07320.
XX
PR 29-JUN-1993; 93US-0086428.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Bukh J, Miller RH, Purcell RH;
XX
DR WPI; 1995-061006/08.
DR P-PSDB; AAR69681.
XX
XX
PT Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
XX
PS Claim 1; Pages 77-78; 186pp; English.
XX
CC AAQ83892 encodes AAR69681 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate SA6, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines, antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 576 BP; 101 A; 156 C; 171 G; 148 T; 0 other;

Query Match 54.1%; Score 340.8; DB 16; Length 576;
Best Local Similarity 89.7%; Pred. No. 5.2e-96;
Matches 366; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATGCCTCTGGGATTATCATGTTACCAATGATGCGCCAAACTCTTCC 279
DB 1 GTTCCTTACCGAATGCCTCTGGGTTGATCATGTACCAATGATGCGCCAAACTCTTCC 60
QY 280 ATAGTCTATGAGCAGATAACCTGATCTACAGCACCTGGTTGCGTTCGTGTGTCATG 339
```

Db	61	ATAGTCTATGAGGCTGATGACCTGATCCCTACACGACCTGGCTGGCTGCCCTGTGTC	120
Qy	340	ACAGGTAATGTAGTAGATGCTGGGTCCAAATTACCCCTACACTGTCTAGCCGCCGAGCCTC	399
Db	121	AAGGATAATGTACGTAGATGCTGGGTTTCATACCCCCACACTATCAGCCCCGAGCCTC	180
Qy	400	GGAGCAGTCAACGGCTCCTCTTCGGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC	459
Db	181	GGAGCGGTCAACGGCTCCTCTTCGGAGGGCCGTTGATTACTTGGCGGAGGGGCGCCCTG	240
Qy	460	TGCTCCGCGTTATACGTAGAGACGCGTGTGGGGCACTATTCTTGTAGGCCCAATGTTTC	519
Db	241	TGCTCCGCGTTATACGTTCGGAGACGTGTGGGGGCATTGTTTGTAGGCCCAATGTTTC	300
Qy	520	ACCTATAGGCTCCGACGACGCTACGCTGCAGAACTGCAACTGTTTCCATTTACAGTGGC	579
Db	301	ACCTATAGGCTCCGACGACGCTACGCTGCAGAACTGCAACTGTTTCCATTTACAGTGGC	360
Qy	580	CATGTTACCGGCCACCGGATGGCATGGATATGATGAACCTGGTAA	627
Db	361	CATATCACTGGCCACCGGATGGCATGGACATGATGAATTGGTCA	408

Search completed: December 19, 2003, 18:51:24
Job time : 175.169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:03:34 ; Search time 1620.71 Seconds
(without alignments)
9447.586 Million cell updates/sec

Title: US-09-899-303a-31
Perfect score: 630
Sequence: 1 ATGGTAGGTCATCGATAC.....TCGATGACTGCGTAATAG 630

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gsa1:*

29: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	57.6	9.1	488	9 AV755731	AV755731
C 2	39.6	6.3	492	9 AV758366	AV758366
C 3	38.4	6.1	920	29 CNS0062R	AL061710 Drosophila
C 4	37.8	6.0	1002	13 BX374637	BX374637

C 5	37	5.9	944	13	BX437004
C 6	37	5.9	1201	13	BX443048
C 7	36.6	5.8	703	10	BF479339
C 8	36.6	5.8	721	12	BI462364
C 9	36.4	5.8	962	29	CNS0140F
C 10	36.4	5.8	1201	13	BX356664
C 11	35.8	5.7	918	13	BX384351
C 12	35.4	5.6	405	9	AA021112
C 13	35.4	5.6	962	10	BG571230
C 14	35.4	5.6	1178	13	BX397483
C 15	35.2	5.6	727	13	BU425620
C 16	35.2	5.6	791	29	CNS00AKB
C 17	35	5.6	632	10	BG297765
C 18	35	5.6	903	28	AY531723
C 19	34.8	5.5	1101	29	CNS002BV
C 20	34.6	5.5	715	29	AG031744
C 21	34.6	5.5	879	10	BG198380
C 22	34.4	5.5	448	12	BI478979
C 23	34.4	5.5	678	12	BM332179
C 24	34.4	5.5	771	12	BM348565
C 25	34.4	5.5	947	12	BI838659
C 26	34.2	5.4	229	10	AW932968
C 27	34.2	5.4	359	13	BX370241
C 28	34.2	5.4	458	28	BH121594
C 29	34.2	5.4	534	12	BJ082331
C 30	34.2	5.4	761	29	CNS020DP
C 31	34.2	5.4	1201	13	BX335532
C 32	34.2	5.4	1201	13	BX376097
C 33	34	5.4	560	9	AA607428
C 34	34	5.4	733	14	CD445238
C 35	33.8	5.4	690	29	CC315210
C 36	33.8	5.4	884	29	CNS02Y9J
C 37	33.8	5.4	1201	13	BX387694
C 38	33.6	5.3	331	10	BE040504
C 39	33.6	5.3	449	14	CA500407
C 40	33.6	5.3	682	28	BH316210
C 41	33.6	5.3	784	12	BG920015
C 42	33.6	5.3	847	14	CA987143
C 43	33.6	5.3	903	12	BG855135
C 44	33.6	5.3	1101	29	CNS002SC
C 45	33.6	5.3	1176	13	BU184099

ALIGNMENTS

RESULT 1
AV755731/c 488 bp mRNA linear EST 19-OCT-2000
LOCUS AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
DEFINITION AV755731
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 488)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.

Unpublished
Homo sapiens cDNA BM clones

CONTACT: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES

```

source
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN

Query Match 9.1%; Score 57.6; DB 9; Length 488;
Best Local Similarity 66.3%; Pred. No. 9.4e-05;
Matches 114; Conservative 0; Mismatches 54; Indels 4; Gaps 2;

QY 455 CCTCTCTCCGGTTATAGTAGGAGACCGGTGTGGGACATTTCTTGGTAGGCCAAA 514
Db 459 CGCTCTCTCAGCTCTCTACGTGTGGGACCTCTGCGACGGAGTGATGCTTGCAGTTTCAGC 400

QY 515 TGTTCACTATAGCTCGCAGCAGCAGCTACGCTGAGAGTCTCAACTGTTCATTATCA 574
Db 399 TGATCATCT---GGCTTCAGCAGCAGCTAGTTGTGCGATGATGCAACTGCTCCATCTATC 343

QY 575 GTGGCCATGTTTACCGGCCACCGGATG-GCATGGGATATGATGATGAAGTGGT 625
Db 342 CTGGGCCATCACTGGACACCGTATGAGCATGGGACATGATGATGAAGTGGT 291

RESULT 2
AV758366/c 492 bp mRNA linear EST 19-OCT-2000
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
Liu,X., Qian,B., Liu,P., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN

Query Match 6.3%; Score 39.6; DB 9; Length 492;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 102; Conservative 0; Mismatches 64; Indels 4; Gaps 2;

QY 457 CTCTGCTCCGGTTATAGTAGGAGACCGGTGTGGGACATTTCTTGGTAGGCCAAA 516
Db 458 CTGTGATCAGCTCACTACGTTGTGGGACCTCTCGTTGGGTTATCGTTGACGCCAATG 399

QY 517 TTCACTTATAGGCTCGCCAGCAGCGTAGCGGTGAGAGTCAACTGTTCATTATCA 576
Db 398 ATTATCTCT---CAGCAGCAACATTTGTTGTGCAAGATGCAACTGCTCATTTATCCT 342

QY 577 GGCCATGTTACCGG-CCACCGGATGCGATGGGATATGATGATGAAGTGGT 625
Db 341 GGCTGCATCACTGCACTACAGTATGCGCATAGGCTATGATGATGAAGTGGT 292

CNS0062R 920 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR13J24 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL061710 GI:4943910
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 920)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1. .920
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR13J24"
/clone_lib="RPCI-98"
/notes="end : TET3"
BASE COUNT 307 a 62 c 39 g 122 t 390 others
ORIGIN

Query Match 6.1%; Score 38.4; DB 29; Length 920;
Best Local Similarity 18.0%; Pred. No. 15;
Matches 55; Conservative 124; Mismatches 126; Indels 0; Gaps 0;

QY 93 GGCTCTCGCACCGGTGTAGGCTCTTGGAGCGGGTAAACTATGCAACAGGGAATTT 152
Db 416 KBBKKKCAANBDBCYRKRSKTGAVKYSGBKRBGKAYKKSGBSRDGAKBBDASGH 475

QY 153 ACCCGGTGTCTTCTCTATCTTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 212
Db 476 DKRSTCYAVCKENANABYSBCTSKYBTSYGVAAATTTAAATTTKYAATCNCYMC 535

QY 213 CTCTGCACTTCCCTACCGAATGCTCTGGATTATCATGTTATACCAATGATGCCCAA 272

```


[illegible]

Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
 Plate: LLAM11680 row: k column: 05
 High quality sequence stop: 712.
 Location/Qualifiers
 1. .721
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5269756"
 /lab_host="DH108"
 /clone_lib="NIH_MGC_97"
 /notes="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTT-TTT-TTT-TTT-TTN-3',
 size-selected for average insert size 2.2 kb and
 normalized to Rf 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 175 a 170 c 209 g 167 t
 ORIGIN

Query Match 5.8%; Score 36.6; DB 12; Length 721;
 Best Local Similarity 55.0%; Pred. No. 44;
 Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 373 ACCCTTACTGTACGCCGAGCGCTCGGAGCAGTCAAGCTCTCTTCGAGAGCGGTT 432
 |||||
 Db 240 ACCACTAACTGTTAGCCCGCCATCTCTCCCGCGCGCGCTCGTCTCAAGCGCT 181
 |||||

QY 433 GACTACTAGCGAGGCGTCCCTGCTGCTCGCTTATACAGAGCGGTGGG 492
 |||||
 Db 180 CGCTCTCCGCGCGCGGCGCTCTACCCCGCGGCGCTTCAGGGGCGCTGAGCCGG 121
 |||||

QY 493 GCACTATTCTT 503
 |||||
 Db 120 GCACTAAGCGT 110
 |||||

RESULT 9
 CNS0140F/c
 LOCUS
 DEFINITION
 Drosophila melanogaster genome survey sequence 17 end of BAC
 BACN12U12 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL104409
 AL104409.1 GI:5616020
 GSS.
 Drosophila melanogaster (fruit fly)
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 962)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage :
 Submitted (23-JUL-1999) Genoscope - FRANCE (E-mail : seqref@genoscope.cns.fr
 BP 191 91006 EVRY cedex - FRANCE)
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBeloBAC11.

FEATURES
 source
 Location/Qualifiers
 1. .962
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN12U12"
 /clone_lib="DrosBAC"
 /plasmid="pBeloBAC11"
 /notes="end : T7"
 BASE COUNT 226 a 156 c 169 g 207 t 204 others
 ORIGIN

Query Match 5.8%; Score 36.4; DB 29; Length 962;
 Best Local Similarity 35.3%; Pred. No. 52;
 Matches 110; Conservative 40; Mismatches 160; Indels 0; Gaps 0;

QY 231 AAATGCTCTGGGATTATCATGTTACCAATGATTCGCCAAATCTTCTCCATAGTATGA 290
 |||||
 Db 648 AATGGCGGCTHACATDTTGTTCGCCGACATTAATCTCAWGTCTCCYAAARAAAAA 589
 |||||

QY 291 GGAGATAACCTGATCTACACGACCTGGTTCGGTGCCTTGTGTGTCATGACAGTAATGT 350
 |||||
 Db 588 MAAKAAATAATGAARCTCTGBCAATCACATTGSCCATTTTCCCATGGRVRRKR 529
 |||||

QY 351 GAGTAGATGCTGGGTCAAATTAACCCCTACACTGTGAGCCCCAGCCTCGGAGAGTCAC 410
 |||||
 Db 528 TRWCCGKKGATGGCACCAMVMDKCGGTAAACMSGATCAGCTGTCTACAGCGATCCSGCTY 469
 |||||

QY 411 GGCCTCTCTTCGGAGAGCGGTTTGACTACCTAGCGGGGCGGCTGCTCTCCCGGTT 470
 |||||
 Db 468 TAGKCCGCCAGTGGAGGAGGAGKGYTKTGTGTTTGGAAAGGTATTGCTCGGGGAC 409
 |||||

QY 471 ATACGTAGGAGAGCGCTGTGGGCACTATTCTTGGTAGGCAAAATGTTACCTATAGGCC 530
 |||||
 Db 408 TGGCNCGGGKATGTTGCKKBTGKCCCTCGTTTGTGCTGCCACSCCCYCNNTTCCC 349
 |||||

QY 531 TCGCCAGCAC 540
 |||||
 Db 348 CCCCHCCWC 339
 |||||

RESULT 10
 BX356664
 LOCUS
 DEFINITION
 BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1015YB03 3-PRIME, mRNA sequence.
 BX356664.1 GI:30378083
 EST.
 ACCESSION
 BX356664
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequençage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1015CA02NP1.
 Location/Qualifiers
 1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1015YB03"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

FEATURES
 source
 Location/Qualifiers
 1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1015YB03"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." 252 c 194 g 261 t 131 others

Query Match	5.7%	Score 35.8;	DB 13;	Length 918;
Best Local Similarity	36.0%;	Pred. No. 74;		
Matches	72;	Conservative	35;	Mismatches 93;
			Indels	0;
			Gaps	0
168	CTCTATCTTTATTCTTGCTCTTCTCTCGGCTGTGACCGTTCCGGCCTCTGCAGTTCCTTA	227		
Db	CTTTTWTYGTGACACMGCTGTMRCMWTTTCWACAGWCTACTCTCTMTYCTTTCTGTGCGCCT	742		
583				
228	CCGAATACGCTCTGGGATTTATCATGTATTACCAATGATTGCCAAACTCTTCCCATAGTCTA	287		
Qy				
Db	CTAMWTGGMTGTMGTGTTTSTGAGCTTWTGMAITGKGGCTTWGUKWMTTCAARAACTA	802		
743				
288	TGAGCGAGATAACCTGATCCTTACACGACACCTGGTTCGGTGCCTTGTGTCATGACAGGTAA	347		
Qy				
Db	803	WAAAAAANAANKNBSSITYMTTSTTTCGCTTTCGGCTTCCTCCYCKGCGSGCTWAKCTAV	862	
348	TGTGATGATAGTGTGGGTCC	367		
Qy				
Db	863	CCKGTCCTTGGTCTMTGKCY	882	

RESULT 12	AA021112	LOCUS	AA021112	405 bp	linear	EST 30-JAN-199
DEFINITION	ze65a12.r1 Soares retina N2b4HR				Homo sapiens cDNA clone	
	IMAGE:363a38				5' mRNA sequence.	

SOURCE: Homo sapiens (human); Homo sapiens (Neanderthal); Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 405)

REFERENCE
Hillier, L.G., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chisoso, S., Dietrich, N., DuBucque, T., Favell, A., Gish, W., Hawkins, B., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlstrand, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9): 807-828 (1996)

```

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 391.
Location/Qualifiers
1. .405
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1280541"
/db_xref="taxon:9606"
/clone="IMAGE:363838"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"

```


QY	167	TCTCTATCTTATCTTCTCTCTCGTCTGACCGTTTCGGCTCTCGAGTTCCT	226
		: : : : : : : : :	
Db	1044	TKACKTKCTBKKBKAKMTKCTTKMBACAMATKQMBTKKTKTCCCTKNTBACCMTK	985
QY	227	ACCGAAATGCCTCTGGGATTTATCATGTTACCAATGATGCCAAACTCTTCCATAGTCT	286
		: : : : : : : : :	
Db	984	BMCCCYKKKCKWNHCYCMTKTMTCCCTTMMCTTKTTMCMVMCMCTKTYCKMKMCCMT	925
QY	287	ATGAGCGCAGATAACCTGATCTCTACACGCACTGGTTCG	324
		: : : : : : : : :	
Db	924	HTKCMCCYCMKMHMCCKCCCBMTKTKMKMCCCMKKKC	887

RESULT	15
BU425620	
LOCUS	727 bp mRNA linear EST 29-NOV-2002 603961926f1 CSEQRN09 Gallus gallus cDNA clone CHEST938d8 5', mRNA
DEFINITION	sequence.
ACCESSION	BU425620
VERSION	BU425620.1 GI:25918296
KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus
.	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus. 1 (bases 1 to 727) Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J. A Comprehensive Collection Of Chicken CDNA Curr. Biol. 12 (22), 1965-1969 (2002)
REFERENCE	
AUTHORS	Contact: Simon Hubbard
TITLE	Department of Biomolecular Sciences
JOURNAL	University of Manchester Institute of Science and Technology (UMIST)
MEDLINE)
PUBMED	PO Box 88, Manchester, M60 1QD, UK
COMMENT	Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk.

/clone_size_lengths)
 /notes=vector: pBluescript II KS(+); Site 1: EcoRI;
 Site2: NotI; This normalized library was constructed from
 1 million independent clones. cDNA synthesis was initiated
 using an oligo(dT) primer, using methylated C in the first
 strand synthesis reaction. Following this first strand
 reaction, double-stranded cDNA was blunted, ligated to
 NotI adapters, digested with EcoRI, size-selected, and
 cloned into the NotI and EcoRI compatible sites of a
 custom modified MCS of the pBluescript (KS+) vector. The
 library was normalized in 2 rounds using conditions
 adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
 Bonaldo et al., Genome Research 6 (1996): 791, except that
 a significantly longer reannealing hybridization was
 used."

Query Match

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2003, 18:11:23 ; Search time 45.4354 Seconds
(without alignments)
6120.154 Million cell updates/sec

Title: US-09-899-303A-31
Perfect score: 630
Sequence: 1 ATGGGTAGTCTCATGATAC.....TGATGATGAACGTGTAATAG 630

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	100.0	630	3	US-08-612-973-31
2	630	100.0	630	3	US-08-927-597-31
3	347.2	55.1	576	2	US-08-086-428B-47
4	347.2	55.1	576	2	US-08-468-570-47
5	347.2	55.1	576	2	US-08-290-665A-47
6	347.2	55.1	576	4	US-08-466-601A-47
7	347.2	55.1	576	5	PCT-US95-10398-47
8	344	54.6	576	1	US-08-086-428B-49
9	344	54.6	576	2	US-08-468-570-49
10	344	54.6	576	2	US-08-290-665A-49
11	344	54.6	576	4	US-08-466-601A-49
12	344	54.6	576	5	PCT-US95-10398-49
13	340.8	54.1	576	1	US-08-086-428B-48
14	340.8	54.1	576	2	US-08-468-570-48
15	340.8	54.1	576	2	US-08-290-665A-48
16	340.8	54.1	576	4	PCT-US95-10398-48
17	340.8	54.1	576	5	PCT-US95-10398-48
18	337.6	53.6	576	1	US-08-086-428B-46
19	337.6	53.6	576	2	US-08-468-570-46
20	337.6	53.6	576	2	US-08-290-665A-46
21	337.6	53.6	576	2	US-08-468-570-50
22	337.6	53.6	576	2	US-08-290-665A-46
23	337.6	53.6	576	2	US-08-290-665A-50
24	337.6	53.6	576	4	US-08-466-601A-46
25	337.6	53.6	576	4	US-08-466-601A-50
26	337.6	53.6	576	5	PCT-US95-10398-46
27	337.6	53.6	576	5	PCT-US95-10398-50

28 335.6 53.3 630 3 US-08-612-973-29 Sequence 29, Appl
29 335.6 53.3 630 3 US-08-927-597-29 Sequence 29, Appl
30 331.6 52.6 9595 3 US-09-014-416-4 Sequence 4, Appl
31 331.6 52.6 9599 3 US-09-014-416-6 Sequence 6, Appl
32 331.2 52.6 576 1 US-08-086-428B-45 Sequence 45, Appl
33 331.2 52.6 576 2 US-08-468-570-45 Sequence 45, Appl
34 331.2 52.6 576 2 US-08-290-665A-45 Sequence 45, Appl
35 331.2 52.6 576 4 US-08-466-601A-45 Sequence 45, Appl
36 331.2 52.6 576 5 PCT-US95-10398-45 Sequence 45, Appl
37 325.2 51.6 1539 2 US-08-470-426B-17 Sequence 17, Appl
38 325.2 51.6 1863 2 US-08-470-426B-14 Sequence 14, Appl
39 324.8 51.6 795 3 US-08-612-973-5 Sequence 5, Appl
40 324.8 51.6 795 3 US-08-927-597-5 Sequence 5, Appl
41 324.8 51.6 2082 3 US-08-612-973-47 Sequence 47, Appl
42 324.8 51.6 2082 3 US-08-927-597-47 Sequence 47, Appl
43 324.8 51.6 2433 3 US-08-612-973-49 Sequence 49, Appl
44 324.8 51.6 2433 3 US-08-927-597-49 Sequence 49, Appl
45 324.8 51.6 9472 4 US-08-150-204E-96 Sequence 96, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-31
; Sequence 31, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..627
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..624

```

US-08-612-973-31
Query Match      100.0%; Score 630; DB 3; Length 630;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTAAGGTCATGATACCTACCGTGGCGGATTCGCCGATCTCATGGGATATATCCCG 60
DB 1 ATGGGTAAGGTCATGATACCTACCGTGGCGGATTCGCCGATCTCATGGGATATATCCCG 60

QY 61 CTCGTAGGGGGCCCCATTCGGGGCGCTCGCAAGGGCTTCGCACACGGTGTGAGGGTCCTT 120
DB 61 CTCGTAGGGGGCCCCATTCGGGGCGCTCGCAAGGGCTTCGCACACGGTGTGAGGGTCCTT 120

QY 121 GAGGACGGGGTAACTATGCAACAGGGAATTTACCGGTTGCTCTTCTCTATCTTTATT 180
DB 121 GAGGACGGGGTAACTATGCAACAGGGAATTTACCGGTTGCTCTTCTCTATCTTTATT 180

QY 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGGGCTCTCGAGTTCCTACCGAAATGCCTCT 240
DB 181 CTTGCTCTTCTCTCGTGTCTGACCGTTCGGGCTCTCGAGTTCCTACCGAAATGCCTCT 240

QY 241 GGGATTATATGTTACCAATGATTCGCCAACTCTTCCATAGTCTATGAGGCAGATAAC 300
DB 241 GGGATTATATGTTACCAATGATTCGCCAACTCTTCCATAGTCTATGAGGCAGATAAC 300

QY 301 CTGATCCTACACGCACCTGGTTGCTGTGTGTCATGACAGGTAATGTGAGTAGATGC 360
DB 301 CTGATCCTACACGCACCTGGTTGCTGTGTGTCATGACAGGTAATGTGAGTAGATGC 360

QY 361 TGGGTCCTAAATACCCCTACCTGTCAGCCCGGAGCTTCGAGCAGTTCACGGCTCCTCTT 420
DB 361 TGGGTCCTAAATACCCCTACCTGTCAGCCCGGAGCTTCGAGCAGTTCACGGCTCCTCTT 420

QY 421 CGGAGAGCGGTTGACTACCTAGCGGGAGGGCTGCCCTCTGTCGGGTTATACGTAGGA 480
DB 421 CGGAGAGCGGTTGACTACCTAGCGGGAGGGCTGCCCTCTGTCGGGTTATACGTAGGA 480

QY 481 GACGCGTGTGGGGCACTATTCTTTGGTAGGCGCAAAATGTTCACTATAGGCTCGCCAGC 540
DB 481 GACGCGTGTGGGGCACTATTCTTTGGTAGGCGCAAAATGTTCACTATAGGCTCGCCAGC 540

RESULT 2
US-08-927-597-31
; Sequence 31, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P. C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```


QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATGCGCAAACTCTTCC 279
Db 1 GTCCCTACCGAATGCTCTGGGTTTATCATGTACCAATGATGCGCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGACATACCTGATCCTACAGCACCTGTTGCGTGTGTCATG 339
Db 61 ATAGTCTATGAGGACATACCTGATCCTGATGTCAGCACCTGTTGCGTGTGTCATG 120
QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATTTACCCCTACACTGTGTCAGCCCGAGCCTC 399
Db 121 GAAGGTAAATGTAGTAGGCTGGGTCCAAATTTACCCCTACACTGTGTCAGCCCGAGCCTC 180
QY 400 GGAGCACTACAGGCTCTCTTTCGGAGAGCGGTGATCTACCTAGCGGAGGCGCTGCCCTC 459
Db 181 GGAGCGGTACAGGCTCTCTTTCGGAGAGCGGTGATCTACCTAGCGGAGGCGCTGCCCTC 240
QY 460 TGCTCCGCTTATACGTAGGACGCGTGGGGCACTATTCTGTTAGGCCAAATGTTTC 519
Db 241 TGCTCCGCACTATACGTGCGGAGCGGTGCGGGCGAGTGTCTTGTAGGCCAAATGTTTC 300
QY 520 ACCTATAGGCTCGCCAGCACGCTACGGTGCAGAACTGCAACTGTTTCCATTTACAGTGGC 579
Db 301 ACCTATAGGCTCGCCAGCATACTACGGTGCAGAACTGCAACTGTTTCCATTTACAGCGGC 360
QY 580 CATGTTACCGCCACCGATGGCATGGGATATGATGATGAATGGTAA 627
Db 361 CATATCACCGCCACCGAATGGCATGGGACATGATGATGAATGGTCA 408

RESULT 5

US-08-290-665A-47
; Sequence 47, Application US/08290665A
; Patent No. 5882852

GENERAL INFORMATION:

APPLICANT: BUKH, J., MILLER, R.H. AND

TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED

TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND

TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS

TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES

NUMBER OF SEQUENCES: 263

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,665A

FILING DATE: 15-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: RICHARD W. BORK

REGISTRATION NUMBER: 36,459

REFERENCE/DOCKET NUMBER: 2026-4116

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 576 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

ORIGINAL SOURCE:

ORGANISM: hom sapiens
INDIVIDUAL ISOLATE: SAS
US-08-290-665A-47
Query Match 55.1%; Score 347.2; DB 2; Length 576;
Best Local Similarity 90.7%; Pred. No. 1.5e-101;
Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTTACCAATGATGCGCAAACTCTTCC 279
Db 1 GTCCCTACCGAATGCTCTGGGTTTATCATGTACCAATGATGCGCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGACATACCTGATCCTACAGCACCTGTTGCGTGTGTCATG 339
Db 61 ATAGTCTATGAGGACATACCTGATGTCAGCACCTGTTGCGTGTGTCATG 120
QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATTTACCCCTACACTGTGTCAGCCCGAGCCTC 399
Db 121 GAAGGTAAATGTAGTAGGCTGGGTCCAAATTTACCCCTACACTGTGTCAGCCCGAGCCTC 180
QY 400 GGAGCACTACAGGCTCTCTTTCGGAGAGCGGTGATCTACCTAGCGGAGGCGCTGCCCTC 459
Db 181 GGAGCGGTACAGGCTCTCTTTCGGAGAGCGGTGATCTACCTAGCGGAGGCGCTGCCCTC 240
QY 460 TGCTCCGCTTATACGTAGGACGCGTGGGGCACTATTCTGTTAGGCCAAATGTTTC 519
Db 241 TGCTCCGCACTATACGTGCGGAGCGGTGCGGGCGAGTGTCTTGTAGGCCAAATGTTTC 300
QY 520 ACCTATAGGCTCGCCAGCACGCTACGGTGCAGAACTGCAACTGTTTCCATTTACAGTGGC 579
Db 301 ACCTATAGGCTCGCCAGCATACTACGGTGCAGAACTGCAACTGTTTCCATTTACAGCGGC 360
QY 580 CATGTTACCGCCACCGATGGCATGGGATATGATGATGAATGGTAA 627
Db 361 CATATCACCGCCACCGAATGGCATGGGACATGATGATGAATGGTCA 408

RESULT 6

US-08-466-601A-47
; Sequence 47, Application US/08466601A
; Patent No. 6572864

GENERAL INFORMATION:

APPLICANT: BUKH, J., MILLER, R.H. AND

TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED

TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE

TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE

TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN

TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES

NUMBER OF SEQUENCES: 160

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,601A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/086,428

FILING DATE: 29-JUN-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: RICHARD W. BORK

REGISTRATION NUMBER: 36,459

```
; REFERENCE/DOCKET NUMBER: 2026-4070US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SAS
; US-08-466-601A-47

Query Match 55.1%; Score 347.2; DB 4; Length 576;
Best Local Similarity 90.7%; Pred. No. 1.5e-101;
Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 220 GTTCCCTACGGAATGCTCTGGGATTATCATGTTACCAATGATGCCCCAAACTCTTCC 279
Db |||||
1 GTCCCTACCGAATGCTCTGGGTTTATCATGTCAACCAATGATGCCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGCAGACATACCTGATCCCTACAGCACCTGGTTGCGTCTGTGTCATG 339
Db |||||
61 ATAGTCTACGAGGCTGATACCTGATCTGACGACCTGGTTGCGTCTGTGTCATG 120

QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 399
Db |||||
121 GAAGGTAATGTAGTAGGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 180

QY 400 GGAGCAGTCAAGGCTCTCTTCGAGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 459
Db |||||
181 GGAGCAGTCAAGGCTCTCTTCGAGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 240

QY 460 TGCTCCGCTTATACGTAGGAGACGCTGTGGGCACTATTTCTTGGTAGCCAAATGTTTC 519
Db |||||
241 TGCTCCGCTTATACGTAGGAGACGCTGTGGGCACTATTTCTTGGTAGCCAAATGTTTC 300

QY 520 ACCTATAGGCTCGCCAGCAGCTACGCTGAGGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 579
Db |||||
301 ACCTATAGGCTCGCCAGCAGCTACGCTGAGGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 360

QY 580 CATGTTACCGGCCACCGATGGCATGGATATGATGATGAATGTTAA 627
Db |||||
361 CATATCAGCGCCACCGATGGCATGGATATGATGATGAATGTTAA 408

RESULT 7
PCT-US95-10398-47
; Sequence 47, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: SAS
; PCT-US95-10398-47

Query Match 55.1%; Score 347.2; DB 5; Length 576;
Best Local Similarity 90.7%; Pred. No. 1.5e-101;
Matches 370; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 220 GTTCCCTACGGAATGCTCTGGGATTATCATGTTACCAATGATGCCCCAAACTCTTCC 279
Db |||||
1 GTCCCTACCGAATGCTCTGGGTTTATCATGTCAACCAATGATGCCCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGCAGACATACCTGATCCCTACAGCACCTGGTTGCGTCTGTGTCATG 339
Db |||||
61 ATAGTCTACGAGGCTGATACCTGATCTGACGACCTGGTTGCGTCTGTGTCATG 120

QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 399
Db |||||
121 GAAGGTAATGTAGTAGGCTGGGTCCAAATACCCCTACACTGTGAGCCCGAGCCTC 180

QY 400 GGAGCAGTCAAGGCTCTCTTCGAGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 459
Db |||||
181 GGAGCAGTCAAGGCTCTCTTCGAGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 240

QY 460 TGCTCCGCTTATACGTAGGAGACGCTGTGGGCACTATTTCTTGGTAGCCAAATGTTTC 519
Db |||||
241 TGCTCCGCTTATACGTAGGAGACGCTGTGGGCACTATTTCTTGGTAGCCAAATGTTTC 300

QY 520 ACCTATAGGCTCGCCAGCAGCTACGCTGAGGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 579
Db |||||
301 ACCTATAGGCTCGCCAGCAGCTACGCTGAGGAGCGGTGAGCAGCCTGGTTGCGTCTGTGTCATG 360

QY 580 CATGTTACCGGCCACCGATGGCATGGATATGATGATGAATGTTAA 627
Db |||||
361 CATATCAGCGCCACCGATGGCATGGATATGATGATGAATGTTAA 408

RESULT 8
US-08-086-428B-49
; Sequence 49, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
```

;; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
;; NUMBER OF SEQUENCES: 159
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & FINNEGAN
;; STREET: 345 PARK AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/086,428B
;; FILING DATE: 29-JUN-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: RICHARD W. BORK
;; REGISTRATION NUMBER: 36,459
;; REFERENCE/DOCKET NUMBER: 2026-4070
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 751-6849
;; TELEFAX: (212) 751-6849
;; TELEX: 421792
;; INFORMATION FOR SEQ ID NO: 49:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 576 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; ORGANISM: homosapiens
;; INDIVIDUAL ISOLATE: SA7
;; US-08-086-428B-49

Query Match 54.6%; Score 344; DB 1; Length 576;
Best Local Similarity 90.2%; Pred. No. 1.5e-100;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 220 GTTCCCTACCGAAATGCTCTGGGATTTATCATGTACCAATGATTCGCAAACTCTTCC 279
DB 1 GTCCCTACCGAAATGCTCTGGGATTTATCATGTACCAATGATTCGCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGAGATTAACCTGATCTACACGACCTGGTTCGCTTGTGTGATG 339
DB 61 ATAGTCTATGAGGAGATTAACCTGATCTACACGACCTGGTTCGCTTGTGTGATG 120
QY 400 GAGCAGTCACGGCTCTCTTCGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC 459
DB 181 GGAGCGGTACGGCTCTCTTCGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC 240
QY 460 TGCTCCGGTTATACGTAGGAGACCGGTGTGGGGCACTATTCTTGTAGGCCAAATGTT 519
DB 241 TGCTCCGGCTATACGTGGGAGACCGGTGTGGGGCACTATTCTTGTAGGCCAAATGTT 300
QY 520 ACCTATAGGCTCGGACGACCTAGGTGCGAGACTGCACTGTTCCATTTACAGTGGC 579
DB 301 AGCTATAGGCTCGGACGACCTAGGTGCGAGACTGCACTGTTCCATTTACAGTGGC 360
QY 580 CATGTTACCGGACCGGATGGATGGATATGATGATGATGATGATGATGATGATGATG 627
DB 361 CATATACCGGACCGGATGGATGGATGATGATGATGATGATGATGATGATGATGATG 408

RESULT 9
US-08-468-570-49

; Sequence 49, Application US/08468570

;; Patent No. 5871962
;; GENERAL INFORMATION:
;; APPLICANT: BUKH, J., MILLER, R.H. AND
;; APPLICANT: PURCELL, R.H.
;; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
;; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
;; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
;; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
;; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
;; NUMBER OF SEQUENCES: 159
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & FINNEGAN
;; STREET: 345 PARK AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/468,570
;; FILING DATE: 6-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/086,428
;; FILING DATE: 29-JUN-1993
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: RICHARD W. BORK
;; REGISTRATION NUMBER: 36,459
;; REFERENCE/DOCKET NUMBER: 2026-4070US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 751-6849
;; TELEFAX: (212) 751-6849
;; TELEX: 421792
;; INFORMATION FOR SEQ ID NO: 49:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 576 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; ORGANISM: homosapiens
;; INDIVIDUAL ISOLATE: SA7
;; US-08-468-570-49

Query Match 54.6%; Score 344; DB 2; Length 576;
Best Local Similarity 90.2%; Pred. No. 1.5e-100;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 220 GTTCCCTACCGAAATGCTCTGGGATTTATCATGTACCAATGATTCGCAAACTCTTCC 279
DB 1 GTCCCTACCGAAATGCTCTGGGATTTATCATGTACCAATGATTCGCAAACTCTTCC 60
QY 280 ATAGTCTATGAGGAGATTAACCTGATCTACACGACCTGGTTCGCTTGTGTGATG 339
DB 61 ATAGTCTATGAGGAGATTAACCTGATCTACACGACCTGGTTCGCTTGTGTGATG 120
QY 340 ACAGGTAATGTAGTAGATGCTGGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 399
DB 121 CAAATAATGTAGTAGGCTGGGTCCAAATACCCCTACACTGTACGCCCCGAGCCTC 180
QY 400 GGAGCAGTCACGGCTCTCTTCGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC 459
DB 181 GGAGCGGTACGGCTCTCTTCGAGAGCCGTTGACTACCTAGCGGAGGGGCTGCCCTC 240
QY 460 TGCTCCGGTTATACGTAGGAGACCGGTGTGGGGCACTATTCTTGTAGGCCAAATGTT 519
DB 241 TGCTCCGGCTATACGTGGGAGACCGGTGTGGGGCACTATTCTTGTAGGCCAAATGTT 300

QY 520 ACCTATAGGCTGCCAGCAGCTACGGTGCAGAACTGCAACTGTTCCATTATACAGTGGC 579
Db 301 AGTATAGGCTGCCAGCAGCACTACGGTGCAGAACTGTTCCATTATACAGTGGC 360
QY 580 CATGTTACCGGGCCACCGGATGGCATGGATGATGATGAATGCTAA 627
Db 361 CATATACCGGGCCACCGAATGGCATGGATGATGATGATGCTCA 408

RESULT 10
US-08-290-665A-49
; Sequence 49, Application US/08290665A
; Patent No. 5892852
; GENERAL INFORMATION:
; APPLICANT: BURK, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORGANISM: hom sapiens
; INDIVIDUAL ISOLATE: SA7
US-08-290-665A-49

Query Match 54.6%; Score 344; DB 2; Length 576;
Best Local Similarity 90.2%; Pred. No. 1.5e-100;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAAATGCTCGGATTTATCATGTTACCAATGATGTCGCAAACTCTTCC 279
Db 1 GTCCCTACCGAAATGCTCGGATTTATCATGTTACCAATGATGTCGCGCACTCTTCC 60
QY 280 ATAGTCTATGAGGAGATACCTGATCCTACGCACTGTTGCGCTGTGTCATG 339
Db 61 ATAGTCTATGAGGAGATACCTGATCCTGACGCACTGTTGCGCTGTGTCATG 120
QY 340 ACAGGTAATGATGATGCTGCTGCTCAATACCTACACTGTCAGCCCCGAGCTC 399
Db 121 CAAATAATGATGATGATGCTGCTGCTCAATACCTACACTGTCAGCCCCGAGCTC 180

QY 400 GGAGCAGTACGGCTCTCTTTCGGAGAGCCGTTGATCTACCTAGCGGAGGGGCTGCCCTC 459
Db 181 GGAGCAGTACGGCTCTCTTTCGGAGAGCCGTTGATCTACCTAGCGGAGGGGCTGCCCTC 240
QY 460 TGCTCCGCGTTATACGTAGGAGACGCGTGTGGGGCACTATTCTTGGTAGCCAAATGCTC 519
Db 241 TGCTCCGCGTTATACGTAGGAGACGCGTGTGGGGCACTATTCTTGGTAGCCAAATGCTC 300
QY 520 ACCTATAGGCTGCCAGCAGCTACGGTGCAGAACTGCAACTGTTCCATTATACAGTGGC 579
Db 301 AGTATAGGCTGCCAGCAGCACTACGGTGCAGAACTGTTCCATTATACAGTGGC 360
QY 580 CATGTTACCGGGCCACCGGATGGCATGGATGATGATGAATGCTAA 627
Db 361 CATATACCGGGCCACCGAATGGCATGGATGATGATGATGCTCA 408

RESULT 11
US-08-466-601A-49
; Sequence 49, Application US/08466601A
; Patent No. 6572864
; GENERAL INFORMATION:
; APPLICANT: BURK, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,601A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; ORGANISM: hom sapiens
; INDIVIDUAL ISOLATE: SA7
US-08-466-601A-49

Query Match 54.6%; Score 344; DB 4; Length 576;
Best Local Similarity 90.2%; Pred. No. 1.5e-100;
Matches 368; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

RESULT 13
US-08-086-428B-48
; Sequence 48, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 G
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE US
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK

REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: hominids
INDIVIDUAL ISOLATE: SA6
US-08-086-428B-48

Query Match 54.1%; Score 340.8; DB 1; Length 576;
Best Local Similarity 89.7%; Pred. No. 1.6e-99;
Matches 366; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTACCAATGATTGCCAAACTCTTCC 279
DB 1 GTTCTTACCGAATGCTCTGGGATTTATCATGTACCAATGATTGCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGGAGATACCTGATCTACACGACCTGCTGCGTGGCTGTGTGTCATG 339
DB 61 ATAGTCTATGAGGAGATACCTGATCTACACGACCTGCTGCGTGGCTGTGTGTCATG 120

QY 340 ACAGTAAATGTAGTACGATGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 459
DB 121 AAGGATAATGTAGTACGATGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 180

QY 400 GGAGCAGTACGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 459
DB 181 GGAGCAGTACGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 240

QY 460 TGCTCCGCGTTATACGTTAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 519
DB 241 TGCTCCGCGTTATACGTTAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 300

QY 520 ACCTATAGGCGTCCGAGCAGCTGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 579
DB 301 ACCTATAGGCGTCCGAGCAGCTGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 360

QY 580 CATGTATACCGGCGGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 627
DB 361 CATATCACTGCGGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 408

RESULT 14
US-08-468-570-48
Sequence 48, Application US/08468570
Patent No. 5871962
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:
ADDRESS: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,570
FILING DATE: 6-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086,428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: hominids
INDIVIDUAL ISOLATE: SA6
US-08-468-570-48

Query Match 54.1%; Score 340.8; DB 2; Length 576;
Best Local Similarity 89.7%; Pred. No. 1.6e-99;
Matches 366; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 220 GTTCCCTACCGAATGCTCTGGGATTTATCATGTACCAATGATTGCCAAACTCTTCC 279
DB 1 GTTCTTACCGAATGCTCTGGGATTTATCATGTACCAATGATTGCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGGAGATACCTGATCTACACGACCTGCTGCGTGGCTGTGTGTCATG 339
DB 61 ATAGTCTATGAGGAGATACCTGATCTACACGACCTGCTGCGTGGCTGTGTGTCATG 120

QY 340 ACAGTAAATGTAGTACGATGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 399
DB 121 AAGGATAATGTAGTACGATGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 180

QY 400 GGAGCAGTACGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 459
DB 181 GGAGCAGTACGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 240

QY 460 TGCTCCGCGTTATACGTTAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 519
DB 241 TGCTCCGCGTTATACGTTAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 300

QY 520 ACCTATAGGCGTCCGAGCAGCTGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 579
DB 301 ACCTATAGGCGTCCGAGCAGCTGCTGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 360

QY 580 CATGTATACCGGCGGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 627
DB 361 CATATCACTGCGGAGGAGCGGTGACCTACCTGAGGAGCGGCTGCGCTC 408

RESULT 15
US-08-290-665A-48
Sequence 48, Application US/08290665A
Patent No. 5882852
GENERAL INFORMATION:
APPLICANT: BURKH, J., MILLER, R.H. AND
APPLICANT: PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES

;/ NUMBER OF SEQUENCES: 263
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: MORGAN & FINNEGAN
;/ STREET: 345 PARK AVENUE
;/ CITY: NEW YORK
;/ STATE: NEW YORK
;/ COUNTRY: USA
;/ ZIP: 10154
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: FLOPPY DISK
;/ COMPUTER: IBM PC COMPATIBLE
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: WORDPERFECT 5.1
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/290,665A
;/ FILING DATE: 15-AUG-1994
;/ CLASSIFICATION: 435
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: RICHARD W. BORK
;/ REGISTRATION NUMBER: 36,459
;/ REFERENCE/DOCKET NUMBER: 2026-4116
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (212) 758-4800
;/ TELEFAX: (212) 751-6849
;/ TELEX: 421792
;/ INFORMATION FOR SEQ ID NO: 48:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 576 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ ORIGINAL SOURCE:
;/ ORGANISM: homosapiens
;/ INDIVIDUAL ISOLATE: SA6
;/ US-08-290-665A-48

Query Match 54.1%; Score 340.8; DB 2; Length 576;
Best Local Similarity 89.7%; Pred. No. 1.6e-99;
Matches 366; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 220 GTTCCTACCGAATGCCTCTGGGATTTCATGTTACCAATGATTCGCCAAACTCTTCC 279
Db 1 GTTCCTACCGAATGCCTCTGGGATTTCATGTTACCAATGATTCGCCAAACTCTTCC 60

QY 280 ATAGTCTATGAGGCGAGTAACCTGATCCTACACGACCTGGTTCGTCCTTGTGTCATG 339
Db 61 ATAGTCTATGAGGCGAGTACCTGATCCTACACGACCTGGTTCGTCCTTGTGTCATG 120

QY 340 ACAGGTATGTGAGTACGTTGGTCCAAATTCACCTACACTGTGAGCCCGAGCCTC 399
Db 121 AAGGATATGTGAGTACGTTGGTTCATATCACCCCCACACTATCAGCCCGAGCCTC 180

QY 400 GGAGCAGTTCACGGTCTCTTCGGAGAGCGCTTACCTACCTAGCGGAGGGGCTGCCCTC 459
Db 181 GGAGCGTTCACGGTCTCTTCGGAGAGCGCTTACCTAGCGGAGGGGCGCCCTG 240

QY 460 TGCTCCGGTTATACGTAGAGACGCGTGTGGGCACTATCTTGTGAGCCAAATGTC 519
Db 241 TGCTCCGGTTATACGTAGAGACGCGTGTGGGCACTATCTTGTGAGCCAAATGTC 300

QY 520 ACCTATAGGCTCCGACGACGCTACCGTGCAGACCTGCACTGTTCCATTACAGTGC 579
Db 301 ACCTATAGGCTCCGACGACGCTACCGTGCAGACCTGCACTGTTCCATTACAGTGC 360

QY 580 CATGTTACCGGCCACCGGATGGCATGGGATGATGATGAACCTGGTAA 627
Db 361 CATATCACTGGCCACCGGATGGCATGGGATGATGATGAATGGTCA 408